Reps and representations:
a warm-up to a grammar of lifting[*]

Maria Esipova (University of Oslo)

December 28, 2021

Abstract   In this paper, I outline a grammar of lifting (i.e., resistance training) and compare it to that of language. I approach lifting as a system of generating complex meaning–form correspondences from regularized elements and describe the levels of mental representations and relationships between them that are involved in full command of this system. To be able to do so, I adopt a goal-based conception of meaning, which allows us to talk about mappings from complex goals to complex surface outputs in systems of intentional action, signaling and non-signaling, interactive and non-interactive, in a unified way, and show how it applies in lifting. I then proceed to argue that the grammar of lifting is architecturally very similar to that of language. First, I show that both involve stable (idiomatized/lexicalized) pairings of regularized forms with regularized meanings. Second, I argue that in both lifting and language, meaning–form mapping is mediated by syntax, which, crucially, operates on non-linearized hierarchical structures of abstract objects that include both content morphemes and functional morphemes. I conclude, following and expanding on some insights from prior literature and offering further evidence for them, that neither of these architectural phenomena (idiomatized meaning–form pairings and abstract syntax) is specific to language, with both of them likely emerging in skilled action that does not necessarily involve social interaction, due to considerations of repeatability and reusability of elements in new contexts.

Contents

---

1. Introduction

1.1. A grammar of lifting: why and how?

The "super-linguistic" research program seeks to expand the traditional object of study in linguistics in two broad ways: (i) including historically overlooked meaning-bearing aspects of spoken and written communication, such as hand gestures, facial expressions, prosodic modulations, emoji, etc. into our formal models of linguistic behavior (Lascarides & Stone 2009; Ebert & Ebert 2014; Ebert 2017; Schlenker 2018a,b; Esipova 2019a,b, 2021a; Hunter 2019, a.o.), and (ii) applying the toolkit and the mindset of a linguist to objects that are not human language, namely, animal communication (see, e.g., Coye et al. 2017; Schlenker et al. 2017 for an overview), as well as non-linguistic systems of structured outputs in humans, such as pictures (Abusch 2012, 2019; Rooth & Abusch 2019; Greenberg 2018, 2019; Maier 2019; Maier & Bimpikou 2019; Cohn 2020; Esipova 2021b, a.o.), music (Lerdahl & Jackendoff 1983; Katz & Pesetsky 2011; Schlenker 2019, a.o.), dance (Patel-Grosz et al. 2018; Charnavel 2019; Napoli & Liapis 2019, a.o.), yoga (Hess & Napoli 2008), and knitting (Fruehwald 2016).

    The value of (i) for linguistics is incontestable, since these aspects of communication integrate into our utterances at all levels of representation in systematic ways that affect both their form and their meaning and simply cannot be ignored by linguists seeking empirical adequacy in their work. In other words, there is nothing "super" about this direction of inquiry—it is just plain linguistics.

    The value of (ii), however, is somewhat less obvious, especially when it focuses on behaviors whose primary goals are not communicative in nature. It stands to reason that, say, pictorial communication, e.g., in pictorial narratives or information signs, or interpretative dance can share some properties with linguistic communication and can, thus, be studied using the same tools. But what about other systems of complex patterns of human behavior, such as non-interpretative dance? Even if we focus exclusively on surface outputs in such systems (i.e., their phonology/phonetics), we can still study their systematic properties in comparison to those found in language in order to understand how certain universal tendencies manifest in different systems, as is done, for instance, for syllable structure in yoga

in Hess & Napoli 2008, or for articulatory effort reduction in dance vs. sign in Napoli & Liapis 2019. However, as I will argue in this paper, we can take it beyond the surface form and study the properties of meaning–form mappings in any system that involves systematic correspondences between complex goals ("meaning") and complex surface outputs ("form") in a uniform way.[1] This would further ground the "super-linguistic" research program within cognitive science, allowing us to fully explore its potential to inform our understanding of how the human mind works.

But to do so efficiently, we need to adopt an architecture-driven methodological approach. That is, when trying to outline a grammar of a given system, we should first and foremost aim to explicitly identify the levels of representation involved and how they interact with one another, which has not been done systematically in most work on non-linguistic behavior in humans cited above (with some exceptions, such as Cohn 2020). For instance, research on abstract hierarchical structures in music and dance, i.e., their "syntax", gleaned through observable surface structures, i.e., their "phonology" (Lerdahl & Jackendoff 1983; Katz & Pesetsky 2011; Charnavel 2019), has so far been conducted independently from research on meaning in music and dance, i.e., their "semantics" (Schlenker 2019; Patel-Grosz et al. 2018), and, crucially, without considering how these three levels of representation fit together architecturally—in stark contrast to how linguists have been studying language.

Thus, in this paper, I use an explicitly architectural approach to outline a grammar of *lifting*, a term that I use to refer to the broad idiom of resistance training,[2] i.e., athletic activity aimed at increasing one's strength, endurance, muscle mass, etc. by creating extra tension in a contracting muscle with the help of one's body weight, free weights, resistance bands, weight machines, etc. I approach lifting as a rule-governed generative system, i.e., a system that creates novel complex outputs from regularized elements. The main goal of the paper is to outline the architecture of grammar of said system, i.e., the levels of mental representations and relationships between them involved in full competence in the idiom of lifting, to compare said architecture to what we have in language, and to discuss the implications of this comparison for our understanding of how different architectural phenomena emerge.

Now, what makes lifting a particularly interesting comparison to language in this respect? On the one hand, lifting is closer to language than many other forms of non-linguistic action in that it relies on frequently repeated and highly formulaic movement patterns, which leads to creation of highly conventionalized meaning–form pairings that can be combined and modified in systematic ways (i.e., once again, with predictable mappings between goals and surface outputs). On the other hand, like with other types of athletic activity, the primary goals of lifting have nothing to do with information transfer or even non-directed signaling; moreover, lifting itself is non-interactive—in contrast, for instance, to combat sports. This allows us to study the universal properties of how our minds organize systems of meaning–

---

[1]While Hess & Napoli (2008) do not directly talk about the meaning of yoga poses, Napoli & Liapis (2019) do attribute the differences in how effort reduction works in performance dance vs. performance sign to the difference in their functions. However, in this paper, I will go beyond such broad, general effects that pragmatic considerations can have on phonetic and phonological processes and will instead focus on much more local phenomena, although I will briefly talk about how broad pragmatic considerations can affect the relative optimality of various surface outputs in lifting in subsection 3.3.

[2]Further more specific sub-idioms include, for instance, powerlifting or Olympic weightlifting.

form mappings regardless of whether they involve information transfer or any kind of social interaction. Finally, the mapping between goals and actions is usually very straight-forward and, thus, easier to model—in contrast, for instance, to other skilled action, such as dance or music, or even some other athletic activities, such as climbing, where this mapping can be much more complex and opaque.

1.2. Structure of this paper

In section 2, I introduce the goal-based conception of meaning as one that can apply uniformly to linguistic and non-linguistic action. I then provide arguments for the existence of meaning distinct from form in lifting, based on the potential for nonsense (both at the level of individual movements and at the level of combination of movements) and the fact that one and the same surface contrast can be semantic or phonological/phonetic in the context of a given movement. Finally, I discuss the process of creating stable associations between phonologized forms and regularized meanings (idiomatization, or lexicalization) in both lifting and language, suggesting that this process is not specific to language and presumably originated in non-linguistic action due to considerations of repeatability.

In section 3, I argue that in lifting, like in language, meaning–form mapping is mediated by syntax, which, crucially, operates on non-linearized hierarchical structures of abstract objects that include both content and functional morphemes. The main arguments for this come from the existence of systematic processes in lifting, such as modification and compounding, in which meaning–form mappings need to be represented in abstract ways, and from the diverse and relative nature of sub-optimality in surface outputs arising from the same syntactic structure. I finish by discussing the emergence of abstract syntax and suggesting that this architectural phenomenon is not specific to language either and likely originated in non-linguistic action, as well, due to considerations of repeatability and reusability. I, thus, provide further evidence to similar claims in prior literature (e.g., Sterelny 2012; Fujita 2017).

While in section 3 I focus exclusively on the structure of a single repetition (hf. *rep*) of a given movement pattern, which I take to be one of the syntactic units in lifting, in section 4, I very briefly discuss some meaningful processes in lifting that happen at levels larger than a single rep and make a few very preliminary observations about prosody in lifting.

Section 5 summarizes the main points of the paper and outlines some directions for future research.

2. Meaning in lifting

2.1. A goal-based conception of meaning

While there is a long-standing tradition in formal semantics to model meaning in natural language in terms of truth conditions, not all meaning is amenable to the truth-conditional analysis even in language. In particular, we can produce linguistic expressions with the goal of expressing our immediate emotions or performing a variety of social functions, e.g., signaling something about your identity or building rapport with the addressee.

Imagine, for instance, that you drop something heavy on your foot and yell *Ouch!* or *Damn!* to let out your frustration. Intuitively, this is very different from you asserting *I am frustrated*. The latter assertion can be true or false and can, thus, be contested, despite its

highly subjective nature. But in the former case, you are not trying to communicate anything that can be true or false (in fact, you are probably not trying to communicate anything to anyone at all in this case, as you likely don't produce this utterance with any addressee in mind, not even yourself). Consequently, this kind of meaning is non-negotiable. Such expressive utterances are, furthermore, *performative*, i.e., the act of producing the utterance itself is crucial for achieving its goals (here, letting out one's emotions). By extension, we can also talk about "performative meanings" as a short-cut for "performatively expressed meanings", or "performative meaning–form mappings", to highlight that the act of producing a given form is crucial for expressing the associated meaning. For items that only carry performative meaning, there is no "use" without "mention"; in particular, there is no way to recover said performative meaning during ellipsis or anaphora resolution.

In addition, truth-conditional and non-truth-conditional meanings—at least performative non-truth-conditional meanings—seem to be treated differently by compositional semantics in that, unlike truth-conditional meanings, non-truth-conditional meanings don't have to interact with their surroundings semantically (Esipova 2021a), thus, warranting distinct formal treatment at this level.[3] However, truth-conditional and non-truth-conditional meanings often coexist within linguistic utterances and even within a single item (e.g., in expressive degree intensifiers like *The movie is damn good*), and the items that carry truth-conditional and non-truth-conditional meanings integrate with each other syntactically and/or phonologically in predictable ways. In other words, in language, truth-conditional and non-truth-conditional meanings come together within a single architecture of grammar, whereby we have systematic mappings between complex meanings and complex surface forms, mediated by syntax.

Thus, however we formalize truth-conditional and non-truth-conditional meaning contributions, we need a way to talk about them in a uniform way, and I propose that the way to do so is in terms of goals. This approach has precedence in literature on non-linguistic action (see, e.g., Pavese 2019 and references therein). Pavese (2019), in particular, draws analogies between non-linguistic action and imperatives. I believe that this approach can be generalized to any kind of linguistic utterances, however, as any utterance, just like any other action, can be associated with a certain goal (or multiple goals). We assert things to communicate our beliefs about the world—and possibly to eventually change our addressee's beliefs. We ask questions to find out something about the world. We utter commands and requests to communicate how we want the world to be—and possibly eventually to bring the world in line with our desires.

Some of these goals are associated with cooperative exchange of information, and, thus, it is quite natural to further model the part of discourse that corresponds to such information exchange as an interactive endeavor that trades in questions under discussion, proposals that can be accepted or rejected, etc., which might still have to rely on the notion of truth. But, as said before, expressive and social meanings, for instance, are not part of this information exchange process, so the notion of truth is not relevant for them. Of course, when someone utters an expression carrying such meanings, an external observer can draw all sorts of

---

[3]Such as, for instance, the expressive semantics in Potts 2007, which is distinct from the analyses that attempt to reduce expressive meaning to some type or other of truth-conditional meaning, such as Potts 2005 or Schlenker 2007. The analysis in Potts 2007 arguably still doesn't fully embrace the key role of producing a given form in performative utterances, but I will not delve into this issue in this paper.

inferences about the person who uttered this expression, which can be true or false—in the same way that we can draw an inference, rightly or wrongly, that it is raining outside when someone comes in in wet clothes—but this doesn't mean that we should be modeling the meaning of the expression uttered in terms of these inferences.[4]

Of course, more can and should be said about the goal-based conception of meaning in language. Among other things, we could think about if/how we want to talk about goal-based meaning in language compositionally. For instance, we can associate linguistic expressions that comprise utterances that partake in information exchange with goals of evoking concepts, constructing more complex concepts with more complex linguistic expressions and eventually building up to assertable content. This seems to go in the direction of the non-truth-based framework of meaning composition pursued in Pietroski 2018, but exploring this connection further is beyond the scope of this paper. The main relevant conclusion so far is that a goal-based approach is a more universal approach to meaning than a truth-based approach both within and outside language.

Now, what kind of goals/meanings do we have in lifting? The exact nature of goals in lifting will vary depending on the size of the structure we are looking at. As I said at the end of the Introduction, in this paper, I will primarily focus on goals associated with a single rep, although larger level goals will become relevant when we talk about relative optimality of surface outputs in lifting, and I will further briefly touch upon macro-level goals in subsection 4.1. Within a single rep, the primary goals will typically be to overload a certain muscle or muscle group in a certain way. For instance, the goal of a single rep of a standard bicep curl is to create tension in the bicep muscle through its full range of motion (ROM) while it's shortening (the goal of the *concentric phase* of the rep, which is typically harder) and lengthening (the goal of the *eccentric phase* of the rep, which is typically easier)—a stimulus that is intended to build up through structured repetition over time and eventually lead to increase in strength and/or hypertrophy in the target muscle.

An anonymous reviewer raised the question whether these goals need to be mentally represented by agents. I will not be able to do proper justice to this question in this paper, but I would like to add a brief note. First, we should separate idealized competence in a given system, be it a specific language, lifting, or any other similar system, from what representations an individual agent might have. As far as the latter is concerned, there is, of course, gradience in how aware one might be of the meaning of their actions. Consider, for instance, an actor who produces sentences in a language they do not speak as part of playing a role. They are repeating the form (and can even become very good at it) without actually understanding the meaning—obviously, in this case, they wouldn't have all the same mental representations as in the case of the same sentences being generated by an actual speaker of the language. Similarly, an agent can go through the motions in lifting without understanding the meaning of what they are doing (e.g., simply because they were instructed by a coach, who, hopefully, does know what the meanings of the different movements are)—in this case they do not have the true complete competence that I am trying to model. The difference

---

[4]Nor does it mean that we should ignore them. The disconnect between the meaning intended by the speaker and the conversational effects a given utterance can have, including the various inferences drawn by different external observers, can be drastic (e.g., in the case of slurs), and variation in such effects can be furthermore intentionally exploited by an aware speaker (e.g., in the case of dogwhistles)—thus, such effects should absolutely be modeled. We should just be very clear about what we are modeling.

is that such lack of true complete competence is more common in lifting than in language. That shouldn't prevent us from analyzing the properties of the complete system itself.

Note, however, that even "naïve" lifters, with no explicit instruction in principles of resistance training, can have some intuitions about what different movements do and don't do, due to mind–muscle connection, the strength of which can vary across individuals (and can, of course, be increased with training) and muscle groups. This brings us to a very important property of meaning–form mapping in lifting, namely, its highly non-arbitrary nature, which is a distinct property from performativity.[5] As with expressive and social meaning in language, meaning in lifting is performative, i.e., you need to actually produce a given form to achieve the associated goal. However, expression of meaning in lifting is somewhat different from, say, performative expression of affect through conventionalized items in language in which the form–meaning link is completely arbitrary. The difference is that in the case of lifting, the goals of a given movement can be achieved regardless of whether or not the agent is aware of the link, if the movement is performed properly: e.g., if you properly perform a single rep of a quad-loading movement such as the squat, you will, in fact create muscle tension in your quads, regardless of whether you had this goal explicitly in your mind. In the case of a completely arbitrary link, however, the agent— or the external experiencer—needs to know about said link in order for the performative effect to obtain for them. For instance, this knowledge is crucial for the ability of the act of uttering a swear word or flipping the bird to serve as an outlet for one's emotions (note that while you can yell swear words or flip the bird in a particularly brisk manner, the expressive power of these conventionalized forms is divorced from the expressive power of the acts of yelling or performing a brisk motion, the latter being, once again, non-arbitrary). Similarly, performative effects of uttering arbitrary linguistic forms (e.g., slurs) on external experiencers can obtain only if said observers are aware of the link—but in this case the agent doesn't have to be. In this sense, the connection between meaning and form in lifting is more akin to iconic meaning–form pairings in (near-)linguistic communication (conventionalized or not), the acts of producing which are also performative, as these acts themselves create direct sensory experiences in order to depict objects or events.

Note also that the tight isomorphism between performing an act and achieving its goal in lifting starts to disappear at larger levels: e.g., if you increase the volume of quad training in your program with the goal of growing bigger quads, you might fail at achieving this goal (as many factors other than systematically creating muscle tension in the target muscles will affect whether this goal will be achieved). This is similar to how one might swear in an attempt to relieve their frustration and yet feel no relief, or use a certain phonological feature associated with a certain group in an attempt "to belong" and yet fail at actually belonging. Note that this doesn't mean that the meaning is no longer performative: performing certain actions is still crucial for achieving the goals of said actions, but it might not be sufficient.

Now that we have a basic understanding of how the goal-based conception of meaning can be applied both to language and to lifting, I will proceed to providing two arguments that we do indeed have meaning distinct from form within the idiom of lifting and then discussing conventionalization of meaning–form pairings in lifting.

---

[5]Albeit the two are not unrelated; in particular, non-arbitrariness can be a source of performativity.

2.2. Arguments for meaning in lifting

2.2.1. Potential for nonsense

We know that structural well-formedness, syntactic or phonological, is distinct from having a meaning or making sense. Thus, [blɪk] would be a phonotactically licit word of English—unlike *[bnɪk]—but it is not conventionally associated with any meaning.[6] The first stanza of Lewis Carroll's 'Jabberwocky' poem, composed of exclusively nonce roots, but using English grammar and functional morphemes, is an all-time favorite in intro to linguistics classes to showcase that a sentence needs not make sense to be syntactically well-formed (although the nonce words in 'Jabberwocky' do rely heavily on sound symbolism, so they are arguably not completely meaningless). Chomsky's (1956) *Colorless green ideas sleep furiously* is intended to make a similar point: even though it is composed of actual English words and is syntactically well-formed, our physical reality is such that it is hard for us to imagine an event witness that would make this sentence true under its literal interpretation in any world that resembles ours. In other words, given our world knowledge, this sentence makes no sense.[7]

Similarly, we can have movement patterns in lifting that are biomechanically feasible and might even look like possible exercises on the surface, but to anyone who understands the kinesiological principles of resistance training, they make no sense, i.e., these movement patterns do not actually map onto any reasonable muscle overload goals. This might be because the agent has no muscle overload goals in mind whatsoever (which is often the case with mindless moving around with weights in many "workouts" one can find online)—or they might have specific goals in mind, but don't know how to correctly establish mappings between said goals (meaning) and movements that would achieve them (form), i.e., they don't actually "speak the language" of lifting.

An example of the latter is illustrated in (1), which demonstrates the "chest exercise" created by a social media personality and fitness brand owner Vince Sant a.k.a. VShred (this "exercise" was originally intended to be performed on a cable machine; (1) replicates it with a resistance band).[8] While on the surface it might look like a variation of the standard cable/band chest fly, shown in (2), and is intended to achieve similar goals, the force vector created by the cable/band remains parallel to the arm throughout the movement in (1), i.e., no additional resistance is placed on the pectoral muscles, whose job is to move the arms across the body.

(1)   Meaningless chest "exercise"

---

[6]At least, it wasn't originally, when first introduced in Chomsky & Halle 1965. One could argue that it has by now acquired a meaning in linguistics as the prototypical nonce word, e.g., in the context of terms like *the blick test*.

[7]I sometimes hear people say that the 'colorless green ideas' example shows that syntactic well-formedness is distinct from semantic well-formedness, interpretability, or something along these lines. I would like to stress that compositional semantics has no problem interpreting this sentence. In fact, in order for us to know that this sentence makes no sense, we need to first interpret it compositionally.

[8]VShred's original YouTube video presenting this "exercise" has since been taken down, following criticisms from the YouTube fitness community. See, for instance, this takedown by Jeff Cavaliere from Athlean-X for further details on why this "chest exercise" doesn't actually train the chest: https://youtu.be/dAlpe1eIYeM?t=116 (note: Cavaliere never mentions VShred in the video as the author of the "exercise"; instead, the video contains a recognizable parody of VShred performed by Jesse Laico).

(2)   Standard chest fly

One might, thus, say that, as VShred doesn't actually "speak the language", he ended up creating a "blick" of an exercise. Of course, the analogy with the original point of the "blick" example isn't perfect, and for a good reason: the fact that [blɪk] is not a word of English is an accident—it could have been one, while the exercise in (1) can never be a meaningful exercise within the idiom of lifting (or any training idiom for that matter). That is, of course, because, unlike in language, meaning–form pairings in lifting are never arbitrary. A closer analogy would, therefore, be with how likely something is to be a word of English based on its iconic potential. Thus, [blɪk] is arguably a better word in terms of its iconic potential than, say, [ˈdɛrmi], even though the latter is still phonotactially licit; for instance, [blɪk] would be more likely to be used situationally in spontaneous speech to iconically depict an event of a certain kind—as it is simply more likely that a real-life event would evoke a similar sensory experience in the listener as [blɪk] rather than [ˈdɛrmi]. Going back to the example in (1), the problem, thus, doesn't seem to just be that V-Shred doesn't have the conscious knowledge of the principles of resistance training, but also that he doesn't seem to be able to properly tap into his mind–muscle connection (discussed in the previous subsection) to feel how much tension his target muscles are actually experiencing.

Either way, the main point of the example above is that both in language and lifting, we can assess the meaningfulness of individual items. But we can also assess the meaningfulness of combinations of meaningful items. In particular, while compounding meaningful movements that target different muscles/muscle groups is in general a valid and common process in lifting,[9] which I discuss in more detail in subsection 3.2.2, some movement combinations

---

[9]Note that the term *compound movements* is used in lifting to refer to any movement that recruits multiple joints/muscle groups. Here I am talking specifically about combining movements that already exist as independent idiomatized exercises. Not all multi-joint movements are compounds in this sense, although they can still be analyzed as multi-root "words". I come back to this issue in subsection 3.1 when talking about the morphosyntactic complexity of the deadlift.

make more sense than others. For instance, the combination of a dumbbell bicep curl with an overhead press (a.k.a. curl-press, or curl-to-press), shown in (3), is a meaningful combination of two meaningful (and idiomatized) movements that would use similar weights to properly overload the target muscles and would, furthermore, additionally engage the rotator cuff muscles during the transition between the two movements.[10] It is also a functional movement pattern that can have uses in everyday life. In contrast, combining a bicep curl with a squat, which are also both meaningful (and idiomatized), makes no sense from the perspective of resistance training. First, the lower body muscles targeted by the squat can handle much more weight than the biceps, so the lifter, limited by how much they can curl, wouldn't be properly overloading the squat portion of the compound. Furthermore, there is no specific reason to combine these two movements, as the two target muscle groups are neither complementary nor antagonistic, so there is no benefit to training them together within a single rep, nor is there any additional meaning created by the transition between the two or any ostensible practical use for this movement pattern. In other words, by combining a curl and a squat we have created a "colorless green ideas" of lifting: the individual items within the sequence are meaningful (regardless of whether the meaning–form association is arbitrary), but their combination is not.

(3) Sensical compound: curl-press
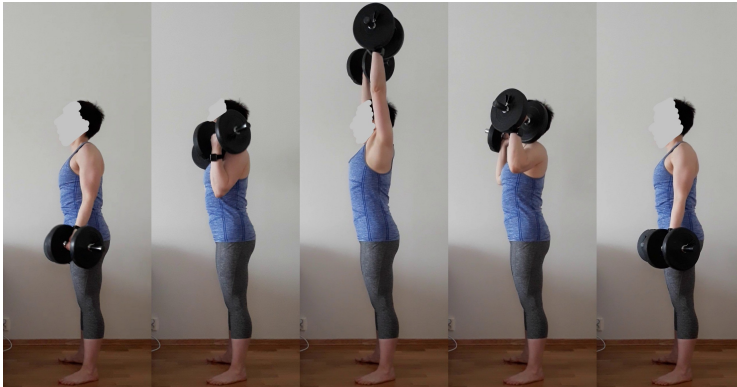
(4) Nonsensical compound: curl-squat

Before I proceed to the next argument for meaning in lifting, let me add a quick note

---

[10]A partial analogy from language would be connecting vowels in compounds like *speed-o-meter*, except in lifting these transitional movements can be adding meaning, while in language they are purely phonological.

clarifying potential confusion. When talking about nonsensical combinations of meaningful elements, I am talking specifically about the idiom of resistance training. In particular, I am not talking about sequences of heterogeneous movements that people can perform as reps within a set without trying to create additional meaning by combining specific movements or optimizing these sequences based on how much weight the various muscle groups involved can handle. One might in principle encounter a combination of a bicep curl and a squat within such a sequence (although they likely wouldn't be linearized as in (4)), but there would be no specific meaning behind putting these two movements next to each other in such a sequence. This is unsurprising, as the overall primary goal of such training can be, for instance, to "burn calories" or to build cardiovascular endurance—which are different from the goals of resistance training (and can, in fact, interfere with them). Now, in some cases, it is possible to combine different types of goals within a single training session or even exercise (e.g., sled-pushing can be done in a more strength-focused vs. more cardio-focused way), but to be able to consciously and effectively balance the different types of goals in this way, one would need to understand the principles of different types of training, which doesn't erase, but, conversely, highlights the fact that different types of training are associated with different types of goals and, thus, constitute different idioms. Consider the following analogy from speech production: we can vocalize with the goal of simply creating noise, which would still be a meaningful action. The vocalizations that we thus produce can be "sensical" sentences in a given language, interpretable but nonsensical sentences, sequences of meaningless syllables, or even vocalizations that do not make use of the regular phonemic inventory of a given language—whatever the case, they can successfully accomplish the goal of creating noise and would, thus, be meaningful from this perspective. And, of course, we could in principle produce vocalizations with both the goal of transferring information and the goal of creating noise; in fact, there is less tension between these two types of goals than in the case of resistance vs. cardio training or "calorie burning". Performing movement sequences that make some, little, or no sense from the perspective of resistance training in order to burn calories (get one's heart rate up, "be active", have fun, etc.) is, thus, similar in this respect to vocalizing in order to create noise.

2.2.2. Semantic vs. phonetic/phonological differences

In language, a certain contrast between two surface structures can be associated with different meanings, or it can be due to phonetic or phonological reasons.

For instance, palm orientation, in general, has a meaning-distinguishing potential in signed languages, i.e., a (categorical) change in palm orientation is a phonemic contrast. For example, the American Sign Language (ASL) signs STARS and SOCKS in (5) are only distinguished by palm orientation, although palm orientation is not itself contributing a piece of meaning, i.e., palm orientation is not morphemic in this case. Palm orientation can be morphemic, however. For example, palm orientation in the ASL possessive pronoun in (6) is morphemic in that it encodes meaningful information about the referent (the palm faces the referent or the locus associated with it). However, a change in palm orientation can also be caused by articulatory considerations (e.g., depending on what other joints are involved when producing a given sign with a smaller or larger amplitude) or phonological processes (e.g., orientation assimilation in compounds or from the non-dominant hand to the dominant hand), and it can be subject to variation across signers (see, e.g., Liddell & Johnson 1989;

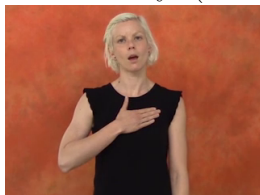Schembri 2001; Sandler & Lillo-Martin 2006; Pfau et al. 2012).

(5) a. STARS (ASL)[11]
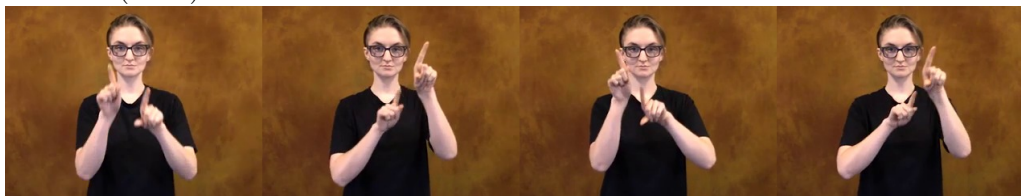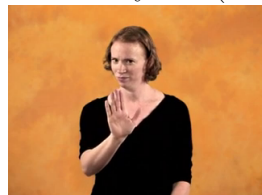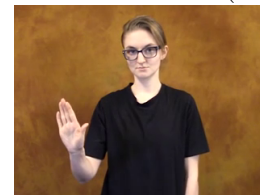

b. SOCKS (ASL)

(6) a. POSS-1 'my' (ASL)   b. POSS-2 'your' (ASL)   c. POSS-a 'their' (ASL)

Similar examples can be given from spoken language. /o/ and /a/ are two distinct phonemes in Russian and can distinguish between two words when they surface in their unreduced form under stress, e.g., /skot/ [skot] 'cattle' vs. /skat/ [skat] 'ray (fish)'. The contrast between the two can also be morphemic, as Russian has monophonemic morphemes that are underlyingly /o/ and /a/, e.g., *šl-a* /ʂla/ [ʂla] 'go.PAST-SG.FEM' vs. *šl-o* /ʂlo/ [ʂlo] 'go.PAST-SG.NEUT'. However, both /o/ and /a/ surface as [a] (or, rather, typically as one of its allophones, [ɐ] or [ʌ]) in certain environments in the process of vowel reduction, e.g., in the syllable immediately preceding the stressed one, as in /poˈʂla/ [pɐˈʂla] 'go.INCH.PAST.SG.FEM'—which is a purely phonological process. Furthermore, the specific patterns of vowel reduction are subject to dialectal variation and can possibly be affected by other social factors.

Similar phenomena can be observed in lifting. For instance, changing the orientation of the grip between pronated and supinated on a pulling, rowing, or curling movement changes the anatomy of the movement in a way that affects the recruitment pattern for the target muscles and is, thus, associated with a change in muscle overload goals. Thus, the standard pull-up, with double-overhand grip, shown in (7a), will load the forearm muscles more, as compared to the underhand grip pull-up (a.k.a. chin-up), with double-underhand grip, shown in (7b), which will load the biceps more. A neutral/hammer grip (not pictured) on the pull-up will engage both muscle groups in a more balanced way. A mixed grip on the pull-up, shown in (7c), will target the two muscles groups asymmetrically (which can be desirable) and will add a further anti-rotational component to the movement. Changing the orientation of the grip also affects the ROM for the other muscles involved in the movement, in particular,

---

[11]Screenshots for ASL signs are from https://spreadthesign.com/

the lats, which are the primary mover in all variations of the movement. Thus, the choice of grip on the pull-up affects its meaning. The grip contrasts on pulling/rowing/curling movements, thus, closely resemble the meaningful surface contrasts in language discussed above, especially the morphemic ones.

(7) a. Overhand grip   b. Underhand grip   c. Mixed grip

In contrast, changing the orientation of the grip on the barbell deadlift, whose compositional structure and various surface properties are discussed in greater detail in the next section (see (10a) for an illustration), between double-overhand and mixed isn't associated with different muscle overload goals, as this doesn't significantly affect the anatomy of the movement pertaining to the primary target muscles (knee and hip extensors). Instead, the choice of grip on the deadlift depends primarily on articulatory considerations.[12] Many lifters use the mixed grip with heavier weights, as it makes it easier to hold on to the barbell, even though it creates an unwanted asymmetry and is less safe (for one, it creates a higher risk of a bicep tear in the supinated arm). In addition, lifters might have existing muscle asymmetries, which can make them choose the specific version of the mixed grip (right overhand/left underhand vs. left overhand/right overhand) that they are strongest with for their heaviest sets. Finally, social considerations can come into play as well, for instance, if a lifter thinks they look "cooler", more attractive, etc. when deadlifting with a specific grip. The grip contrasts on the deadlift, thus, resemble the contrasts due to phonetic or phonological considerations in language discussed above.

2.3. Conventionalization of meaning–form pairings

Note that the pronated–supinated distinction discussed in the previous subsection is gradient, and the muscle recruitment pattern of a given pulling/rowing/curling movement changes continuously in line with the grip change. This is also true for changing the width of the grip on such movements, changing the angle of the bench press to target the different parts of the pectoral muscles, changing the angle of the upper body on a Bulgarian split squat to target the anterior vs. posterior chain more, etc. In other words, in lifting, we often deal with an inherently gradient one-to-one mapping between form and meaning. This, of course, isn't the case for the natural language examples discussed in the previous subsection: phonemic contrasts are perceived as categorical, and so are the contrasts in meaning. Trying to produce

---

[12]The trap bar deadlift, which employs a neutral grip, will target knee and hip extensors differently from the barbell deadlift, but that's due to a different weight distribution, not the change in grip orientation. Similarly, deadlifting with dumbbells will allow both for a more variable grip orientation and for a more variable weight distribution, but the two are independent of one another (beyond some obvious articulatory considerations).

something between /o/ and /a/ in /skVt/ in Russian will be understood as a sloppy instance of either /skot/ 'cattle' or /skat/ 'ray (fish)', not a word denoting a hybrid of a cow and a manta, or a cownose ray.

In this sense, one might think that meaning–form mapping in lifting would be akin to iconic gradient meaning–form mapping in (near-)linguistic communication, such as changes in movement path in classifier predicates in sign or in gesture of non-signers.[13] This would make sense, since in both cases the mapping between meaning and form is non-arbitrary, and non-arbitrariness seems to be a pre-requisite for gradient meaning–form mapping (the reverse doesn't seem to be true in the sense that conventionalized meaning–form pairings relying on categorical distinctions can have a non-arbitrary source, but conventionalization does famously reduce iconicity in language). However, despite the inherently gradient nature of meaning–form mapping in lifting, lifting primarily operates on conventionalized meaning–form pairings, which, thus, rely on categorical distinctions. This applies both to movements themselves (idiomatized exercises and their potentially meaningful parts) and to operations on syntactic units (reps and sets); the latter will be discussed in the next section, here I will focus on the former. For instance, the common idiomatized variations of the dumbbell bicep curl based on the grip orientation include: regular curl, with fully supinated grip; reverse curl, with fully pronated grip; hammer curl, with neutral grip; Arnold curl, going from fully pronated at the bottom to fully supinated at the top. The form of these is conceptualized as categorical, i.e., articulatory differences between individual productions within a given variation are ignored, and the muscle overload goals are similarly formulated in a categorical fashion, e.g., "targeting the inner/outer biceps" or "targeting the forearms" or "loading the bicep throughout both flexion and supination simlutaneously"—even though the correspondence between the surface movement pattern and the actual muscle recruitment pattern remains, of course, perfectly gradient.

This is unsurprising, as resistance training relies on repeating the same movement patterns, and repeatability requires stable articulatory targets, which, of course, then get reinforced with further repetition. Thus, both in language and in lifting, repetition is crucial for

---

[13]See, e.g., Goldin-Meadow & Brentari 2017 for a discussion of categorical and gradient contrasts in sign. Note, however, that Goldin-Meadow & Brentari call gradient aspects of sign "gestural" and seem to furthermore assume they are "non-linguistic". I object to both characterizations. The use of the word "gestural" in this case doesn't match the layperson use of this word, which describes a certain mode of articulation in spoken communication that can be used to encode either gradient or categorical contrasts. This terminological choice, thus, might lead one to believe—mistakenly—that mode of communication directly and categorically prescribes the gradient vs. categorical nature of contrasts in this mode, and is, furthermore, spoken-language-centric. As for the tendency to label gradient meaning–form mapping "non-linguistic", it is, of course, omnipresent in linguistics, going back to Hockett 1959. However, the mere juxtaposition of "linguistic" and "non-linguistic" aspects of communication seems to presuppose the existence of a monolithic language module in human mind. Yet, language is a complex interconnected system that brings together multiple types of representations. We look up some meaning–form pairings in a mental dictionary, however it is organized, and establish others situationally via some iconic module, but the two can come together in a coherent compositional structure within a single utterance (see also a discussion of such integration in subsection 3.1). If anything, the fact that we can create meaning–form pairings on the spot and then know how to integrate them into our utterances in a systematic way syntactically, compositionally, linearly, and prosodically is a testament to how complex and productive our linguistic ability is. On the flipside, as this paper shows, conventionalized meaning–form pairings relying on categorical distinctions are not specific to communication, which is another reason to avoid essentially equating such pairings with language.

the emergence of phonological representations for surface outputs as abstractions over multiple tokens.[14] Regularization of forms, i.e., "phonologization", of course, goes hand in hand with regularization of meanings associated with those forms—once again, in both lifting and language, and for both arbitrary and non-arbitrary pairings. Examples of regularization of form and meaning in the process of conventionalization of originally gradient meaning–form pairings in language involve using a fixed number of repetitions to conventionally denote a plurality of individuals or events in speech, sign, or gesture (see, e.g., Schlenker & Lamberton 2019 on various repetition-based plurals in sign and gesture) or regularization of prosodic high degree modification (see Esipova 2019b on the latter, as compared to simple segment lengthening to indicate increased length or duration). As discussed in subsection 3.4, considerations of repeatability also play a role in the emergence of syntax, once the patterns that need to be repeated become too complex to be efficiently stored as atomic units.

Not only do we need to be able to repeat certain movements in the context of lifting proper, but, as noted in Sterelny 2012 for skilled action in general, we may also need to demonstrate these movements without the intent to achieve the associated muscle tension goals (and, consequently, often with little or no additional load) when teaching them to other individuals.[15] Note that such demonstrations also represent a step away from performativity. An even further step would be then to turn such demonstrations into gestures used to simply evoke the concept of a certain movement or, even further, to related and potentially more abstract concepts (training in general, strength, etc.)—a process that would go hand in hand with simplification of form. This is somewhat similar to the connection between the gestures we can use to ask someone to give us something or to depict someone giving us something (non-performative), properly regularized in signed languages as verbs meaning 'give' or 'take', and actually attempting to take it from them (performative), or between gestures we use to tell someone to move in a certain direction (non-performative) and actually pushing them into that direction (performative).[16] Note, however, that in the case of both taking/pushing or spontaneous gesture depicting taking/pushing, we are dealing with ad hoc, non-regularized meaning–form mapping, which, thus, does not come with the same amount of architectural complexity that we observe in full-blown systems of language or lifting.

Apart from repeatability, another (related) consideration in conventionalization of meaning–form pairings in both language and lifting is contrast, which is needed for establishing categorical distinctions and, arguably, further reinforces their categorical nature. In language, contrast is crucial to be able to reliably distinguish between different meanings or forms, be it for the purposes of conventionalization or when juxtaposing several non-conventionalized meaning–form pairings within a single discourse situation, i.e., when we create a situational categorical distinction. In lifting, contrast also plays a practical role. There is little value in doing three variations of the same movement that only differ minimally, say, flat, 10%

---

[14]Or, alternatively, "clouds" of stored experiences, as in exemplar-based theories (see Kaplan 2017 for a literature overview on the topic).

[15]Sterelny also mentions performing various elements of a given skill without their normal outcome and often without all of their normal substrate for offline practice, e.g., in various music-related skills. As far as I can tell, completely offline practice is not particularly useful in lifting, although, of course, performing exercises with lower loads is used for warm-up.

[16]See also Esipova 2021a for a discussion of other cases of going "from performatives to performances" in (near-)linguistic communication.

incline, and 10% decline bench press, as the differences in the muscle recruitment patterns across these three are negligible—as opposed to, say, flat, 45% incline, and 30% decline.

Before moving on, let me add a quick note about combining conventionalized and non-conventionalized meaning–form mappings. In fn. 13, I have already mentioned that non-conventionalized depictions can be integrated into utterances that are otherwise composed of conventionalized elements (see also section 3.1). However, the reverse is also possible: conventionalized elements can be inserted, deliberately or not, into otherwise non-conventionalized depictions. For instance, ASL signs and handshapes seep into co-speech gesture of fluent L2 signers (e.g., Weisberg et al. 2020). Anecdotally, this also happens with conventionalized movements from lifting (or other athletic activities), which can seep into more ad hoc action. Note that this is not always because this is the most practically efficient movement pattern in a given situation (just like integrating ASL signs or handshapes into gesture is not always the most efficient way to communicate with non-signers): for instance, "deadlifting" or "squatting up" a dropped pen is arguably not the most energy-efficient way to pick it up, yet this is something that at least some experienced lifters would regularly do.[17] Presumably, in cases like this, ease of retrieval is a factor that can favor falling back on more stable representations. A full exploration of this phenomenon and the parallels with sign intrusion in gesture is, however, outside the scope of this paper.

Now that we have seen that movements in lifting do have meaning and have talked about conventionalization of certain meaning–form pairings, let us talk about how we combine and modify meaningful elements in lifting and the architecture underlying these processes.

3. Inverted Y model of grammar of lifting

3.1. Overview

In the generative tradition, it is common to assume the inverted Y model of grammar for language, where meaning–form mapping is mediated by syntax, which operates on non-linearized hierarchical structures. On the right side of the split, compositional semantics interprets the literal meaning of the syntactic structure, and pragmatics builds post-compositional meanings based on the output of compositional semantics, surface form of the utterance, various properties of the context, etc. via further reasoning about the mental states of the speech participants. On the left side of the split, linearization, creation of a prosodic structure, other phonological processes happen (possibly in several passes) to eventually create a pronounceable surface structure. Certain theories, furthermore assume late exponent insertion, whereby syntax operates on abstract objects that have no phonological content, and phonological exponents of specific parts of the syntactic structure get inserted and manipulated post-syntactically during the various stages of syntax→form mapping. It is, furthermore, common to assume that compositional semantics doesn't have direct access to the surface output either (unlike pragmatics). A schematic representation of this model is given in (8).

---

[17]Of course, in the case of heavier objects, this can be the most efficient way to pick them up. However, note that the goals of movements in lifting have to do with creating extra muscle tension rather than moving an object from point A to point B maximally efficiently (in contrast, e.g., to climbing/bouldering), which is why these movements are not always the most biomechanically advantageous way of moving objects. However, as they do often involve moving an object along a certain trajectory, the relevant forms can get activated in tasks that require that.

(8) Inverted Y model of grammar of language
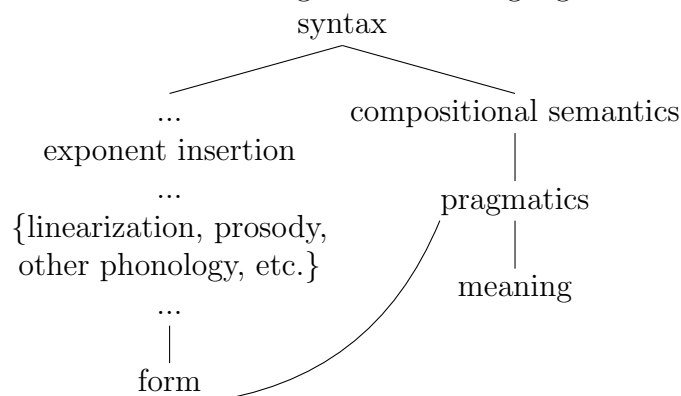
The elementary abstract objects that the syntax operates on are morphemes. It is commonplace to distinguish between two types of morphemes: content morphemes (roots, l-morphemes, etc.), like $\sqrt{\text{CAT}}$ or $\sqrt{\text{BOOM}}$, and functional morphemes (f-morphemes, etc.), like a plural or a past tense morpheme. The details of how various conventionalized meaning–form associations are organized architecturally vary across specific theories in ways that are not directly relevant for this paper. However, for simplicity, I will be making a non-lexicalist assumption that all complex structures are built in the syntax in both language and lifting, regardless of whether they are idiomatized, as in, e.g., Distributed Morphology (Halle & Marantz 1993 et seq.) or Nanosyntax (Starke 2010 et seq.), although I won't be providing any evidence from lifting for this assumption (see also fn. 34).

I will add two notes to this description. First, while meaning–form pairings can be conventionalized within a given language, as mentioned in fn. 13, we can also create ad hoc meaning–form pairings on the spot when communicating, which will not have a stable phonological representation of the form or a properly regularized meaning, but they can nonetheless integrate into linguistic utterances syntactically, compositionally, linearly, and prosodically in predictable ways. For instance, when describing an explosion, we can say *It goes boom*, where *boom* is an ideophone of English: it does have an iconic source, but it has been fully conventionalized. But we could also instead produce a non-conventionalized depiction of an explosion via a non-phonologized vocalization and/or gesture, which would be integrated into the utterance in the same way as *boom*.[18] Second, since syntax and compositional semantics operate on abstract objects, they don't care about the channel of a given exponent, i.e., a given syntactic object can in principle be exponed through any channel (independently of how conventionalized the meaning–form pairing involved is). For instance, in some languages, including English, focus is marked by a pitch accent (suprasegmental marker), while in others, it can be marked by a particle (segmental marker)—but the corresponding syntactic representation would be the same in both cases.[19]
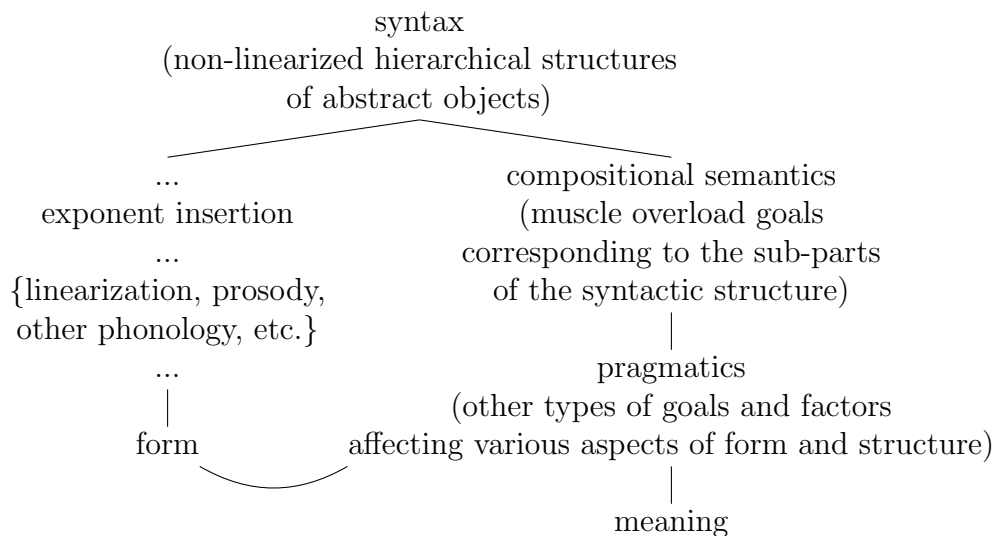
---

[18]See a naturally occurring example of such a depiction here: https://youtu.be/fTvZfbYj7MQ?t=301

[19]In addition, there doesn't have to be a one-to-one mapping between the syntactic structure and the surface structure. For instance, one and the same syntactic object can get multiple surface realizations within one utterance (e.g., focus can be simultaneously marked via a pitch accent and a particle; certain high degree modifiers and affective meanings discussed in Esipova 2019b, 2021a can be simultaneously exponed via prosody and facial expressions, etc.), or even, depending on further assumptions, that pieces of the surface structure can expone overlapping, but non-identical pieces of the syntactic structure. It remains to be seen

In this section, I argue that a similar architecture obtains for lifting as well. In particular, I propose that in lifting, too, meaning–form mapping is mediated by syntax, which, crucially, operates on non-linearized hierarchical structures of abstract objects. These abstract objects can be content/root morphemes, like Elbow Flexion (which is the only content morpheme in the concentric phase of a bicep curl), but they also include modifications such as "1.5-rep" or "cheat rep", discussed in subsection 3.2.1, which are akin to functional morphemes.[20] As we will see, the mapping between meaning and form for such modifications is best represented in an abstract way, because of the variation in possible surface outputs for a single rule, which is the main argument for the proposed architecture. I will also argue that compounding is an abstract syntactic process in lifting, with the syntactic structure, once again, separate from the surface structure. Finally, I will show that the relative optimality of various potential surface realizations of a given syntactic structure can depend both on the biomechanics of a given movement (i.e., phonetic considerations) and on the relative weight of various pragmatic considerations, thus, once again, warranting the separation of the syntactic structure from the surface structure, but allowing pragmatics to affect the latter. Note also that, like language, lifting, too, involves expression of meaning via different channels, for instance, movements (i.e., "segmental material") vs. intentional manipulation of tempo and duration (i.e., "suprasegmental material"; see also subsection 4.2).

A schematic representation of the inverted Y model of grammar of lifting is given in (9).

(9)   Inverted Y model of grammar of lifting

syntax
(non-linearized hierarchical structures
of abstract objects)

...   compositional semantics
exponent insertion   (muscle overload goals
...   corresponding to the sub-parts
{linearization, prosody,   of the syntactic structure)
other phonology, etc.}
...   pragmatics
form   (other types of goals and factors
affecting various aspects of form and structure)

meaning

Before I proceed to discussing the arguments in favor of this model, let me provide an example of what the syntactic structure of a complex movement pattern can look like. In (10b), I give a very simplified syntactic tree for a single rep of the conventional deadlift (or the deadlift *tout court*), whose concentric phase is shown in (10a) (the eccentric phase is just reversing the movement). Here, I only focus on the dynamic components and completely ignore the isometric contractions as well as such meaningful components as stance (conventional vs.

---

if anything like this happens in lifting.

[20]Note that these claims about syntax in language and lifting are distinct from simply saying that in non-linguistic action, a larger task can be broken down into smaller subtasks (as in, e.g., Pavese 2019).

sumo). Note also that here I do not associate the eccentric phase with any meaning. The received wisdom is that the eccentric phase of the conventional deadlift is there to simply get back to the starting position and should be performed as quickly as possible, while maintaining control over the weight. This is not the case for the Romanian deadlift, which is a deadlift variation that does not involve putting the weight back on the ground between reps and does, in fact, aim to overload the eccentric, i.e., create tension in the target muscles when they are lengthening. Since the eccentric phase of the conventional deadlift still needs to be performed in a precise, controlled fashion (for safety and to assure proper set-up for the next rep), I assume that lifters still form proper phonological representations for it.[21]

(10) a. Concentric phase of a single rep of the conventional deadlift
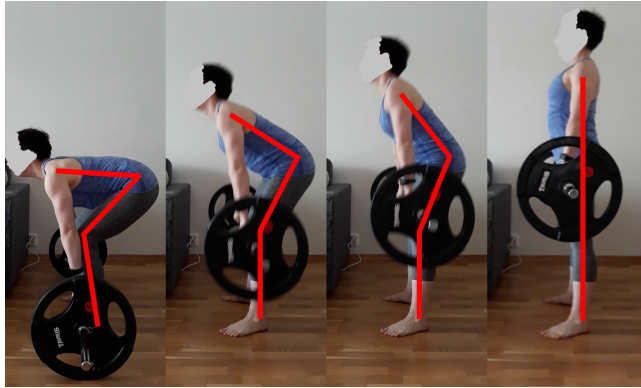
  b.
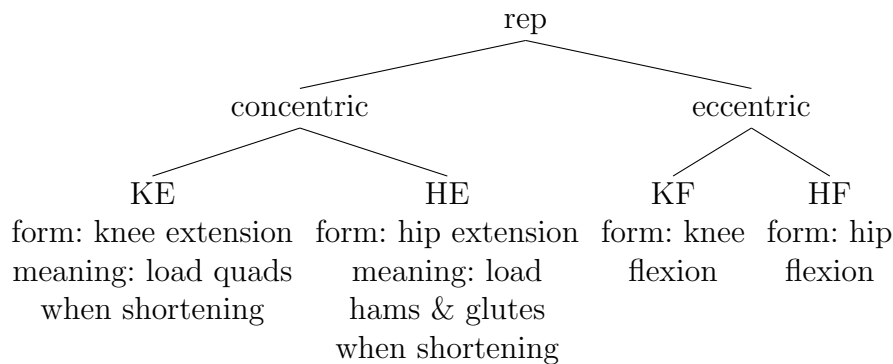
In cases of simple composition like in (10b), the meaning of the mother node is just the conjunction of the goals of its daughters, but we will see instances of more complex compositionality in the next subsection.

Note that here I am treating the concentric phase of the deadlift as a bimorphemic movement, consisting of two meaningful roots, KE and HE, instead of treating KE and HE as two phonological units within a single root. The main justification for this is that there exist monomorphemic variations of the deadlift, the stiff-leg deadlift and the Romanian

---

[21]Of course, it is also possible to completely delete the eccentric phase by simply dropping the weight from the top; similarly, the concentric phase could be deleted, as well, as in eccentric-only pull-ups. I also leave open the possibility that concentric and eccentric phases are not represented syntactically, but are created in the phonology (see also section 4.2). So, e.g., in (10b), we would have a rep simply split into 'knee hinge' and 'hip hinge', with no phase level. Then any modifications that only appear to target one of the phases (e.g., the slow eccentric) would apply at the rep level in the syntax; similarly, deletion of either phase would be handled by positing abstract concentric/eccentric-truncation morphemes in the syntax.

deadlift, that eliminate the knee extension/flexion component[22] and are, thus, explicitly meant to only focus on the posterior chain. In this sense, the deadlift is seemingly distinct from some other multi-joint movements, such as the overhead press. That said, even in the case of those movements, the different components overload different muscles and, thus, have different meaning—so, perhaps, even in those cases, those components should be represented as separate morphemes? This question is, of course, related to the question of how elementary units are identified in action more generally (as asked, e.g., in Pavese 2019 and references therein), but here I ask it in a way that explicitly distinguishes between phonological (i.e., meaning-distinguishing) and morphemic (i.e., meaning-bearing) units. The quandary here, of course, is once again due to the non-arbitrary nature of meaning–form mapping in lifting, but it is by no means specific to non-linguistic action. A similar question arises about ideophones and ideophonic sequences in language, where the the structure of syllables in a sequence and even the structure of a single syllable have been argued to map onto the structure of the event(s) depicted (see, e.g., Thompson & Do 2019 and references therein). For instance, are *ding* and *dong* independent morphemes in *ding-dong*, as they depict two events in a prototypical ringing sequence? Are the onset, the nucleus, and the coda of *ding* independent morphemes, as they depict the different stages of an atomic event in such a sequence? I will leave this quandary unresolved here and will focus on more obvious cases of structures involving multiple morphemes in the subsequent subsections.

Note also that 'knee extension' and 'hip extension' are not very informative specifications of the phonological content inserted into the nodes. For instance, (11) shows the concentric phase of a single rep of the back squat, which has a very similar syntactic structure as in (10b) (assuming that the phases of the back squat are similarly bimorphemic).

(11)   Concentric phase of a single rep of the back squat
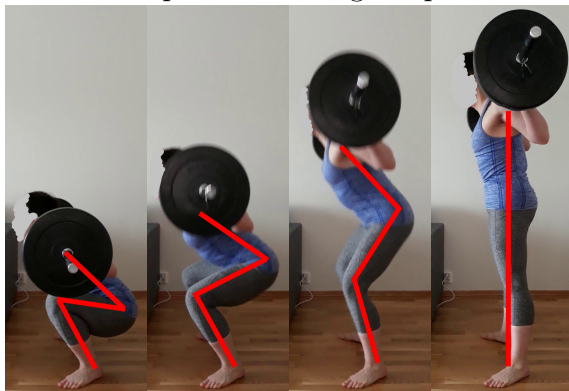
One observable difference between the squat and the deadlift is the relative linearization of the concentric and eccentric phases within a rep: the former precedes the latter on the deadlift; the reverse holds for the squat. Remember, however, that the structure in (10b) is not meant to be linearized, despite what the inevitable 2D nature of tree representations might suggest, i.e., under the proposed model, this is not a syntactic or semantic difference between the two. I come back to issues of linearization in subsection 3.3.1.

---

[22]There might still be some amount of "phonetic" knee extension/flexion on a stiff-leg or a Romanian deadlift, depending on one's body proportions and flexibility, but the assumption here is that this component is eliminated from the syntactic structure.

The other major difference is that, due to the position of the weight and consequent weight distribution, the deadlift involves more of the hip hinge movement and less of the knee hinge movement, as compared to the squat, making it a more posterior-chain-focused movement.[23] In a more elaborate representation, this difference would be captured by specifying the start and the end position of the relevant body parts. I will not attempt to come up with a system of formulating such representations here, but I will note a couple of things in this respect. First, these representations need to be formulated in sufficiently abstract terms, as the exact angles will depend on the lifter's body proportions (and the height of the plates in the case of the deadlift) and will likely have to reference various external objects, such as pieces of equipment. Second, an interesting relevant question is if/how a given lifter's mental phonological representations for various movement patterns are affected by the training cues they used when learning them. The flipside of this question is whether the relative efficiency of training cues for learning movement patterns can help us identify universal tendencies in how humans prefer to mentally represent such patterns. Note that these questions about learning can be asked for L2 learning as well (this analogy could be particularly illuminating in the case of signed languages, with their larger articulators).

Before I proceed to discussing the various phenomena that highlight the need for abstract syntax in lifting, let me add three quick notes. First, just like the model of grammar in language in (8), the one I propose for lifting in (9) is not meant to be a model of production (or processing, for that matter, although in the case of lifting there is less need for a model of processing), but a model of the cognitive architecture that underlies and constrains possible models of production/processing. Second, regarding the "Y-ness" of the inverted Y model: the main point of this paper, the evidence for which is presented in the rest of this section, is that we have syntax as a separate module that builds non-linearized hierarchical structures from abstract morphemes (content and functional) that are devoid of phonological content. I will not present any arguments from lifting that there are any constraints on these structures that are exclusive to syntax and completely blind to anything that happens on the form side or the meaning side. In fact, I have no hard commitments about whether such constraints exist in language, or if all syntax is just Merge (external and internal), with everything else happening at the interfaces. Regardless, the implication of this paper is that both language and lifting make use of the same cognitive module to create syntactic representations, even if said representations in lifting do not make use of the full power of this module.

3.2. Syntactic processes

3.2.1. Modification

While lifting is very formulaic and relies heavily on idiomatized exercises, there is place there for creative processes. New exercises can be created from scratch or as innovative variations of existing exercises. More importantly for our purposes, however, there exist modification patterns that, once learnt, can be productively applied to new cases. Crucially, certain modifications can be applied to different parts of the movement, depending on one's specific goals, and, furthermore, the exact surface outputs resulting from these modifications will

---

[23]And, as mentioned before in fn. 12, the trap bar deadlift will have a weight distribution and subsequent knee/hip hinge ratio that is more similar to that of the squat. Relatedly, the sumo deadlift differs from the conventional deadlift in the lifter's stance, which will also affect the biomechanics of the rest of the movement.

vary depending on the biomechanics of the movement being modified, in a way reminiscent of phonologically conditioned allomorphy in language—all while the underlying principle of the meaning–form mapping involved remains the same. In other words, when learning a new modification, a lifter learns a principle of creating new surface outputs, given a certain abstract hierarchical structure, not multiple unrelated meaning–form pairings. This favors a model of grammar that posits syntax as a mediator between form and meaning and treats such productive modifications as abstract functional morphemes.[24]

Let us look at some examples. One such productive modification is the "1.5-rep" modification, whereby the lifter goes through a certain portion of the ROM twice within a single rep to increase time under tension for the target muscle(s) in that portion of the ROM. The specific portion of the ROM thus targeted will depend on the specific goals of the lifter; e.g., it can be the hardest portion of the ROM, or the easiest (and, thus, systematically underloaded), one in which they want to improve their technique, one that emphasizes a specific muscle (group) in a complex movement, etc. Yet, in all these cases, we have an abstract shared meaning component. Furthermore, where the target portion of the ROM is in the linear structure will depend on the biomechanic properties of the specific movement (and, in some cases, on the individual properties of the specific lifter)—not unlike how the exact form of a partially reduplicated item in a given language depends on the phonological properties of the input item and cannot be described in purely surface terms (e.g., the reduplicant might have to be the lexically stressed syllable of the input; the reduplicant could be truncated to fit a certain weight; the reduplicant can appear in various positions on the surface, depending on further phonotactic constraints of the language, etc.).[25] For instance, (12) shows the 1.5-rep modification aimed at increasing time under tension for the target muscle(s) throughout the hardest portion of the ROM, as applied to the pull-up, the squat, and the bicep curl. Of course, which portion of a given movement is subjectively the hardest may vary across lifters depending on individual weaknesses, injuries, etc., but these portions are generally the hardest ones based on the strength curves for these movements.

(12) a. Pull-up w/1.5-rep modification targeting the hardest portion of the ROM

---

[24]Note that in linguistics it is not uncommon to assume that the inventory of functional morphemes in language is universal and fixed. I do not necessarily subscribe to this view for language, and I most certainly make no such claims for lifting or any other similar system. For the purposes of this paper, by "functional morphemes", I really just mean things that can only modify other things.

[25]See, e.g., Inkelas 2014 for an overview on reduplication. The parallel with (partial) reduplication in language is particularly apt, given that it can be used to encode intensification of the input meaning, which, of course, is a meaning–form pairing with an iconic source.

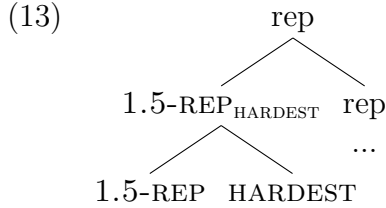b. Squat w/1.5-rep modification targeting the hardest portion of the ROM

c. Bicep curl w/1.5-rep modification targeting the hardest portion of the ROM

The meaning of the modification, i.e., its goal, is the same in all the examples in (12). However, we would have hard time trying to formulate the rule for creating the surface output by only making reference to the linear surface structure, even if we allow ourselves to make reference to concentric vs. eccentric phases of a rep. In (12a), we do the following sequence: 1 concentric–1/2 eccentric–1/2 concentric–1 eccentric; in (12b), we do the reverse: 1 eccentric–1/2 concentric–1/2 eccentric–1 concentric; and in (12c), we reduplicate the mid-range of the ROM. It is implausible that a lifter creates an ever-growing number of unrelated, counter-intuitively formulated atomic meaning–form pairings for all these sub-cases of 1.5-rep modification.[26] Instead, we could capture the regularities across all instances of applying the 1.5-rep modification by positing a syntactic structure in (13), whereby an abstract 1.5-REP morpheme (which is very much like a partial reduplication morpheme in language) combines with a similarly abstract argument that further specifies which portion of the ROM the 1.5-rep modifier will reduplicate (e.g., HARDEST), and the resulting saturated modifier then modifies a rep containing a content morpheme (or several). The result is a new syntactic structure of the rep, and we can then independently create a surface form for this entire structure based on the specific biomechanics of the input movement.

---

[26]This would also predict that someone who learns the 1.5-rep modification for, say, a movement with an ascending strength curve would not be able to apply it correctly to a movement with a descending or bell-shaped strength curve, which could be tested via an equivalent of the "wug test". This is a practical possibility, of course, since, as noted in subsection 2.1, a given individual might be performing movements without understanding their meaning. Once again, here I am talking about idealized competence.

(13)

```
                      rep
              ╱              ╲
      1.5-REP_HARDEST        rep
         ╱      ╲             ...
   1.5-REP    HARDEST
```

An example of a similar, but much simpler productive modification is the "paused rep" modification, where we simply introduce a "pause" into the compositional structure, i.e., an isometric contraction of the target muscle(s), which will also target a specific point of the ROM—typically one's "sticking point", whose surface position will once again depend both on the biomechanic properties of the movement and the lifter.[27]

Yet another example is the "cheat rep" modification, whereby we intentionally use momentum during the concentric phase (e.g., by swinging the weight up on the bicep curl or lateral raise, or by swinging one's chest up to the bar on the pull-up) and then slowly perform the eccentric, which allows overloading the concentric phase explosively, overloading the eccentric with a heavier weight, overloading the target muscle(s) beyond the point of failure, etc. The surface form of the movement, thus, changes quite a bit—the momentum component alone can drastically alter the muscle recruitment patterns of the concentric phase—and it would be, once again, hard and counter-intuitive to try to describe all these changes making reference exclusively to the surface structure.

3.2.2. Compounding

Another syntactic process in lifting I would like to briefly discuss here is compounding, which I already touched upon in subsection 2.2.1. As I said before, this is a process whereby we combine two (or more) existing idiomatized exercises, often in a way that has some added benefits (e.g., recruiting additional muscles, training complementary or antagonistic muscles together, training a functional movement pattern, etc.), i.e., in a way that goes beyond simply conjoining the goals of one movement with the goals of the other. This somewhat resembles how natural language compounds can have idiosyncratic meaning that goes beyond compositionally conjoining the meanings of the two parts, although, of course, the potential for added meaning in lifting is much more limited than in language.

With respect to surface form, similarly to natural language compounds, compounds in lifting aren't just linear sequences of a rep of movement X followed by a rep of movement Y. The exact surface form of a given compound, including the relative linearization of the various sub-components of the movements combined, will, as usual, depend on the biomechanic properties of the specific movements. Thus, the previously discussed curl-press compound, shown earlier in (3), sandwiches a full overhead press rep between the concentric and eccentric phases of a bicep curl rep; plus it also contains a transitional shoulder rotation movement

---

[27]In this case, isometric contractions are meaningful and are, thus, part of the compositional structure. In fact, isometric holds can constitute the entirety of an exercise (e.g., the static plank). Of course, lifters constantly isometrically contract various muscles throughout various movements for stability, safety, and to avoid energy leaks, without these isometric contractions being associated with any primary muscle overload goals—such isometric contractions are then just part of optimal surface outputs. That said, sometimes it is not clear whether or not we should be including isometric contractions into our compositional structure, based on potential secondary muscle overload goals. For instance, the deadlift creates a lot of isometric overload for various muscles of the back—to the point where some consider it a "back exercise".

between the two. Compare it to the squat-push-press, shown in (14), which combines a squat with a push-press, which is itself a modified, more explosive version of an overhead press incorporating an additional leg drive component. There the push-press component follows the squat component, but instead of doing a full squat rep, followed by a full push-press rep, the concentric phase of the squat transitions explosively into the concentric phase of the push-press, incorporating the "push", i.e., the leg drive component of the latter.[28,29]

(14) [Squat-[push-press]]

The model of grammar proposed here for lifting allows us to separate the syntactic process of compounding two movements from the meaning or form idiosyncrasies of specific compounds—just like we can do it for language.

3.3. Syntax→form mapping

Separating syntax from surface form in our model of grammar of lifting also allows us to capture the relative optimality of various surface realizations of one and the same syntactic structure, which can be affected by various considerations, both biomechanic and pragmatic.

3.3.1. Linearization

Unlike in language, there are no ostensible reasons to posit categorical and/or arbitrary rules for linearization of syntactic structures in lifting.[30] Optimal linearization of the various movement sub-components in lifting is instead driven by biomechanic considerations.

For instance, in an optimally produced conventional deadlift, the knee and hip hinge components happen simultaneously, assuring a vertical barbell path over the middle of one's foot and optimizing load distribution throughout the movement, as shown in (15a). Starting

---

[28]On the surface, this is not entirely unlike movement reduction in signed language compounds, whereby two monosyllabic signs combine into a single monosyllabic compound, driven by the general tendency for monosyllabicity in signed languages (see, e.g., Sandler & Lillo-Martin 2006). Of course, in lifting, this process is much less regular, and it is not driven by an independent general tendency for monosyllabicity.

[29]Note that squat-push-press has a much better upper–lower body balance than the hypothesized curl-squat compound in (4), as the squat component is made harder by the need to generate the explosive force necessary for the "push", while the "push" makes the press component easier, so the more force the lifter generates with their legs on the squat, the more weight they can handle on the press.

[30]Arguably, not all linearization in language is categorical or arbitrary either, especially once we start looking beyond segmental material, but parameters like head-initial or head-final seem to be both categorical (for a given type of syntactic structure within a language, not necessarily across the board) and arbitrary.

hip extension before knee extension at the beginning of the concentric phase, before the weight leaves the floor, would be articulatorily impossible, but attempting to do so will likely result in a hip drop, putting the lifter in a disadvantaged position at the start of the movement, and extending at the hips faster than at the knees throughout the movement will disrupt the barbell path and compromise the lifter's balance. These issues are shown in (15b). Extending at the knees too early, as shown in (15c), takes knee extensors out of the picture, increases the moment arm for the hip extensors, and puts more load on the lower back, making it harder to lift the weight and failing at properly fulfilling the goal of loading knee extensors. In other words, doing so unintentionally turns the conventional deadlift into a stiff-leg deadlift, mentioned earlier, but with a vacuous knee extension component (similarly, doing knee extension too early on the squat results in the infamous "good morning squat").

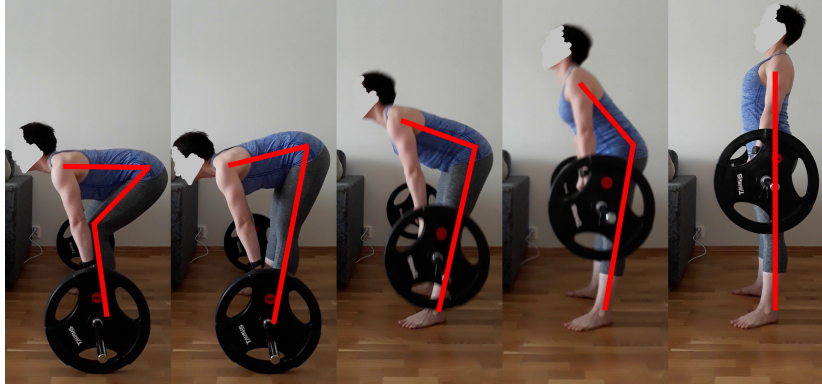(15) Concentric phase of a single rep of the conventional deadlift: linearization

a. Knee extension = hip extension
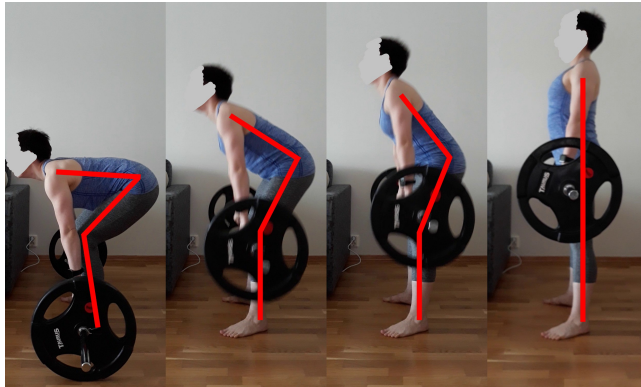
b. Hip extension > knee extension

c. Knee extension > hip extension

Linearization of the concentric vs. eccentric phase within a rep of a given movement is also determined by biomechanic considerations (and oftentimes common sense). For instance, the standard back squat starts on the eccentric rather than the concentric because (i) it is harder to get into position and properly brace one's core at the bottom of the squat, and (ii) starting on the eccentric allows one to use the stretch reflex at the bottom, which helps with getting through the sticking point of the squat. Very similar considerations explain why, for instance, the barbell bench press starts on the eccentric as well. However, both movements can be done from pins, starting on the concentric and bringing the weight to a dead stop on each rep, precisely with the goal of eliminating the stretch reflex and forcing the target muscles to generate all the power necessary to perform the concentric.

3.3.2. Variable relative optimality of surface outputs

While in (15) above, (15a) is incontestably the most optimal candidate, relative optimality of candidate surface outputs in lifting can vary depending on the relative weight a given lifter at a given moment of time places on various relevant considerations, such as lifting as much weight as possible, avoiding injury, avoiding aggravating an already existing injury, looking a certain way when performing the movement, etc. Following an existing tradition in natural language phonology (see, e.g., Pater 2009 for an overview), we could, thus, capture this process of maximizing optimality based on one's priorities via weighted violable constraints, allowing the weights of the constraints to vary. Note that, like in language, this constraint-based approach can be used to analyze both stable phonological representations within a given individual or group (which can nonetheless vary across individuals or groups) as well as potentially variable outputs within a single individual depending on the context.

For instance, in subsection 2.2.2, I have already talked about how the mixed grip on the deadlift creates both biomechanic advantages (ability to lift more weight, as the lifter is less limited by the strength of their grip) and disadvantages (asymmetries, higher risk of injury), as compared to the default double-overhand grip. The choice of grip on the deadlift could, thus, be very crudely modeled as a competition between the 'overhand' and the 'mixed' candidate as evaluated against the two constraints in (16). SAFE penalizes any movement that increases risk of injury relative to some baseline, while AMWAP rewards any movement that allows the lifter to lift more weight relative to some baseline. I will not address the issue of how said baseline is established in any great detail. It can be done, for instance, by treating one of the candidates as unmarked, or by starting with 0 as the baseline for all candidates and assessing how much they increase the risk of injury relative to doing nothing, how much

weight one can lift with this candidate relative to 0, etc. The strategy can in principle vary across constraints. Here I will assume that the 'overhand' candidate is unmarked and will be treating it as the baseline. I will furthermore assume that a constraint can positively or negatively affect the value of a given candidate by adding/subtracting 1 to/from it, multiplied by the constraint's weight coefficient W. Of course, in a more fine-grained model, the candidates will be evaluated against these constraints in a gradient fashion, but for the purposes of our toy derivation, a binary set-up is sufficient: a movement either does trigger a constraint or doesn't. The tableaux in (17) show how either candidate can emerge as the maximally optimal one depending on the relative weights of SAFE and AMWAP.[31]

(16) a. SAFE (Be safe!): $-1*W$ for every movement that increases the risk of injury relative to the baseline
  b. AMWAP (Lift as much weight as possible!): $1*W$ for every movement that allows the lifter to lift more weight relative to the baseline

(17) a.

/deadlift/	SAFE (2)	AMWAP (1)
☞ a. 'overhand' (0)	0	0
b. 'mixed' (-1)	-2	1

  b.

/deadlift/	SAFE (1)	AMWAP (2)
a. 'overhand' (0)	0	0
☞ b. 'mixed' (1)	-1	2

The same toy derivation would work for deadlifting with a neutral back vs. with a rounded (upper) back. The former is in general considered the prescriptively proper form for the deadlift. However, rounding one's back will shorten the moment arm for the hip extensors and will once again allow the lifter to lift more weight, which is why we often see competitive powerlifters lift with a rounded upper back, as shown in (18). Note that experienced powerlifters have stable phonological representations for both potential outputs, i.e., the rounded back output is not a "phonetic accident" for them, like it could be for an inexperienced lifter—which warrants a properly phonological analysis.

(18) Powerlifter Konstantīns Konstantinovs deadlifting with a rounded upper back[32]

---

[31]The weights of the constraints are written in parentheses after their names (no other ordering mechanism is assumed), the value of all the candidates is originally set to 0, the numeric effect of a given candidate being evaluated against a given constraint is given in the appropriate cell of the tableau, and the final value of each candidate after being evaluated against all the constraints is given in parentheses after the candidate's description.

[32]Image source: https://stronglifts.com/deadlift/

3.4. Emergence of abstract syntax

Now that I have provided arguments for the existence of abstract syntax in lifting, let us briefly discuss the implications of this for our ideas about how syntax emerges.

In subsection 2.3, I have already attributed the emergence of phonologized forms paired with regularized meanings to the need for repeatability, in a way that applies both to linguistic and non-linguistic action and presumably originates in the latter. Sterelny (2012) makes similar arguments for the emergence of syntax in skilled non-linguistic action. He argues that both (proto-)language and non-language "involve sequences with structure, and with elements reusable in other contexts" and that some of the pressures for developing syntax as a mediating level of representation (which apply in lifting) involve the need to be able to repeat certain complex patterns with high level of precision and teach them to other individuals (which is easier if they are broken down into smaller parts that can be practiced or demonstrated individually) and the need to be able to re-use various sub-elements of complex action patterns in new contexts—but that the existence of syntax doesn't hinge on the action itself involving social interaction.[33]

Similar claims about syntax emerging in non-linguistic action are made by Fujita (2017), who maintains that "this similarity between language syntax and action syntax reflects a real evolutionary continuity, beyond a mere metaphor". Fujita puts forward a hypothesis of the motor control origin of Merge, whereby "syntax evolved from a preexisting motor control capacity whose original function had nothing to do with language or communication".[34]

I believe that the discussion in this section provides further specific evidence for these claims about the emergence of syntax, and, more specifically, syntax that is non-linearized and operates on abstract objects devoid of phonological content. As the discussion of productive modifications in lifting in subsection 3.2.1 suggests, the latter property of syntax seems to emerge because not only do we need to be able to re-use sub-components of complex actions themselves, but we also need to be able to meaningfully modify different actions in repeatable ways, which must abstract away from the irrelevant variation in the surface properties of the actions being modified and, consequently, the resulting surface outputs. In other words, the abstract nature of syntax is due to the need to create functional morphemes.

4. Further notes

In this section, I will very briefly discuss some of the issues this paper chose not to focus on, but that are nonetheless important for a complete architectural picture.

4.1. Beyond a single rep

So far I have talked exclusively about the structure of a single rep. Of course, reps are integrated into larger structures, as shown in (19).

---

[33]While I support these claims, I would like to make it clear that I do not necessarily subscribe to (or even properly understand) all the aspects of Sterelny's language evolution hypothesis laid out in that paper.

[34]He also argues for the crucial role of Merge in the development of not only syntax, but the lexicon, arguing for a non-lexicalist conception of the lexicon, which doesn't assume a separate "word-building" module. I am sympathetic to this view and am assuming it in this paper, although, as I said in subsection 3.1, I am not offering any evidence from lifting for or against existence of a separate word-building module.

(19)   Basic (non-exhaustive) hierarchy of syntactic and "discourse" units in lifting:
rep > set > exercise session > training session > training mini-cycle (e.g., within a week) > training maxi-cycle (e.g., across several weeks)

Various meaningful and productive processes can happen at all these larger levels as well. Many of them aim to manage the distribution of training volume and intensity for various muscle groups over a given structure. For instance, there are various ways in which working sets (as opposed to warm-up sets, which, of course, also fit into a larger structure in a certain way) can be organized within an exercise session in terms of volume and intensity, e.g., flat, pyramid (start with more reps with lower weights and build up to less reps with higher weights), reverse pyramid, etc. Various specific exercises are organized within and across sessions according to a specific "training split", which specifies which muscle groups are targeted on which days within a mini-cycle ("full-body", "upper-lower", "push-pull-legs", etc.). Finally, training volume and intensity can be managed across maxi-cycles as well.

Some processes happening at these larger levels have more specific goals. For instance, when organizing reps within a set, one productive technique whose purpose is to extend the set beyond the point of failure is the "drop set", where the lifter performs reps with a certain level of intensity to failure and then immediately drops the level of intensity (e.g., by decreasing the weight or switching to an easier variation of the exercise) and performs additional reps to failure. Another productive process, applying to sets, is "supersetting", when single sets of different exercises are performed back-to-back with little or no rest between them, thus, creating a complex set, which can then be repeated. Circuits in circuit training are basically large supersets. As with rep compounding, some supersets are more coherent than others. For instance, it makes sense to superset tricep extensions with bicep curls (antagonistic target muscles, low intensity movements), but it doesn't make much sense to superset heavy squats with heavy deadlifts (similar target muscles, high intensity movements).

It remains to be seen to which extent we can use insights and tools from linguistics, and, in particular, discourse analysis to model such macro-level processes in lifting.

4.2. Prosody in lifting

Another topic I have not touched upon at all is prosody in lifting. There are a few questions one could ask in this respect. What do prosodic structures look like in lifting? Is there an equivalent of prominence marking? To what extent can prosody carry meaning in lifting? I will not aim to properly address any of these questions, but I will offer a few quick ideas that can be developed in the future.[35]

In terms of prosodic hierarchy, the largest potentially relevant unit seems to be the prosodic set (i.e., a unit that maps onto the maximal syntactic set), with smaller prosodic units being prosodic reps (which map onto maximal syntactic reps). Rep boundaries are, in particular, relevant for the 'concentric > eccentric' vs. 'eccentric > concentric' linearization distinction and, in the case of heavy complex movements like the squat, the deadlift, the bench press, etc., for breath reset. Breathing, in general, is much more regularized and consciously controlled in lifting than in speech and, thus, presumably warrants its own tier

---

[35]Here I am only talking about dynamic movements in lifting. Exercises consisting entirely of isometric holds, such as the static plank, L-sit, etc., arguably don't have much of a prosodic structure.

in the phonological structure.[36] It remains to be seen if an intermediate prosodic unit is needed between a prosodic set and a prosodic rep. A prosodic rep, however, needs to be further split into prosodic phases to properly model the processes that can hold between the concentric and the eccentric. It remains to be seen if the full rep and the half-rep within 1.5 reps need to be treated as yet another prosodic phrasing distinction or simply as four phases.

It might be tempting to draw analogies between prosodic structures in lifting and those in signed languages or in spoken languages with highly regularized prosodic phrase structure, such as Bengali in (20) (for instance, peak lengthening and shortening of the target muscle could be treated analogously to L and H tones and assumed to mark phase boundaries). It remains to be seen to what extent such analogies are justified.

(20) Structure of prosodic phrases in Bengali (Khan 2014)
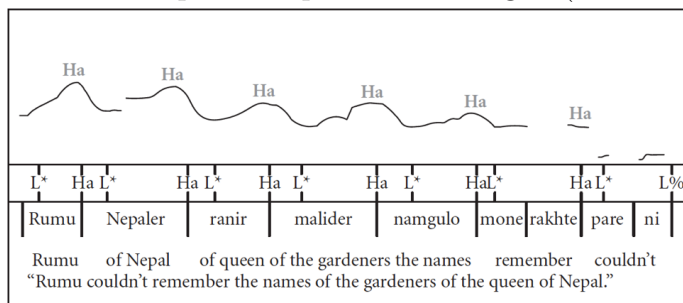
It is, furthermore, unclear to me if there is any equivalent of prominence marking in lifting beyond phrase-edge phenomena. For instance, endpoints of concentric phases that end in a peak contraction of the target muscle(s) are naturally prominent, and this prominence can sometimes be exaggerated, e.g., in an attempt to give the target muscle(s) an "extra squeeze at the top". At this point, it's not clear to me, however, if this should be treated as an instance of prominence marking, nor if other parts of a movement can be systematically accentuated, in particular, in a way that would be associated with a specific meaning.

What definitely does happen in lifting is manipulating tempo and/or duration in a meaningful way—as opposed to changes in tempo and/or duration due to, say, one's level of exhaustion. For instance, "cheat reps", mentioned earlier in 3.2.1, shorten the concentric and lengthen the eccentric. Intentionally slowing down the eccentric is, in general, a common way of increasing the overall time under tension for the target muscle and balancing out the load on the concentric vs. the eccentric. On the other hand, "speed reps", intentionally produced at a high tempo, can be used to load the target muscles more explosively. Just like similar meaningful prosodic modulations in language (such as prosodic degree modification mentioned in subsection 2.3), under the model proposed in (9), modulations of this kind in lifting would be treated as exponents of abstract syntactic objects.

As with macro-level structures, I leave any further exploration of prosody in lifting for future research.

---

[36]Think, for instance, about how the Valsalva maneuver is often used by lifters on certain heavy lifts, such as the squat, to help create and maintain abdominal pressure throughout the lift, whereby the lifter takes a large breath into their belly before starting the rep, keeps trying to exhale against a closed glottis throughout the concentric phase, and releases the air after the point of maximum effort.

5. Conclusion

In this paper, I have attempted to outline a grammar of lifting as a system of generating complex meaning–form correspondences from regularized elements and to compare said grammar to that of language. I have shown that, despite substantial differences in the nature of meaning in lifting and language, the two systems are organized architecturally in a strikingly similar fashion. Using similar arguments as have been used for language, I have argued, among other things, for (i) separation of structural well-formedness and meaningfulness in lifting, (ii) existence of phonologized forms associated with regularized meanings (idiomatized/lexicalized meaning–form pairings), and (iii) existence of syntax as a level operating on non-linearized hierarchical structures of abstract objects (both content and functional morphemes). I concluded that this architecture is not specific to language and likely emerges in skilled action that does not necessarily involve social interaction, due to considerations of repeatability and reusability of different types of elements in new contexts.

I hope that in the future we can test how architecturally similar systems like lifting are to language experimentally. Apart from the equivalent of the "wug test" for productive modifications mentioned in fn. 26, we could also, for instance, measure reactions of experienced lifters to various types of anomalous outputs in lifting (e.g., reps with sub-optimally linearized sub-components; movements that are biomechanically well-formed, but do not seem to map onto any reasonable muscle overload goals; compound reps combining well-executed meaningful movements in a way that doesn't make sense; sloppily executed movements, e.g., with an irregular tempo or an inconsistent range of motion across reps, etc.) and compare them to similar data from language (in particular, from signed languages, which will hopefully eliminate some irrelevant modality effects).

It would also be worthwhile extending the same architecture-driven methodological approach to other non-linguistic systems. As I have already mentioned in the Introduction, I expect interactive athletic activities to be more challenging—but also very interesting—to analyze in this way. For one thing, one would need to embed a grammar that generates surface outputs based on the agent's goals into a model of the participants' reasoning about each other's goals and subsequent behaviors, which would, furthermore, get even more complex once we go from two-participant sports to team sports. Going beyond athletic activities, it would be interesting to see if/how this approach can be applied to systems whose goals concern affecting external objects rather than the agent's body, such as knitting. As a first approximation, the primary goals of specific knitting techniques and modifications have to do with creating objects with specific physical properties, in a way that could be potentially compositional, however, I am leaving a proper exploration of this idea for future research (by someone knowledgeable about knitting).

References

Abusch, Dorit. 2012. Applying discourse semantics and pragmatics to co-reference in picture sequences. In Emmanuel Chemla, Vincent Homer & Grégoire Winterstein (eds.), *Proceedings of Sinn und Bedeutung 17*, 9–25.

Abusch, Dorit. 2019. Possible worlds semantics for pictures. To appear in Matthewson, Lisa, Cécile Meier, Hotze Rullmann, and Thomas Ede Zimmermann (eds.), *The Compan-*

ion to Semantics. Oxford: Wiley. URL https://compling.cis.cornell.edu/mr249/papers/abusch-possible-apr-29-2019.pdf.

Charnavel, Isabelle. 2019. Steps toward a universal grammar of dance: Local grouping structure in basic human movement perception. *Frontiers in Psychology* 10. 1364. doi: 10.3389/fpsyg.2019.01364.

Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on information theory* 2(3). 113–124. doi: 10.1109/TIT.1956.1056813.

Chomsky, Noam & Morris Halle. 1965. Some controversial questions in phonological theory. *Journal of Linguistics* 1(2). 97–138. doi: 10.1017/S0022226700001134.

Cohn, Neil. 2020. Visual narrative comprehension: Universal or not? *Psychonomic Bulletin and Review* 2(27). 266–285. doi: 10.3758/s13423-019-01670-1.

Coye, Camille, Simon Townsend & Alban Lemasson. 2017. From animal communication to linguistics and back: Insight from combinatorial abilities in monkeys and birds. In Louis-Jean Boë, Joël Fagot, Pascal Perrier & Jean-Luc Schwartz (eds.), *Origins of human language: Continuities and discontinuities with nonhuman primates* Speech Production and Perception Vol. 4, 187–232.

Ebert, Cornelia. 2017. Co-speech vs. post-speech gestures. Talk given at *Language and cognition workshop in memory of Peter Bosch*, Osnabrück, February.

Ebert, Cornelia & Christian Ebert. 2014. Gestures, demonstratives, and the attributive/referential distinction. Talk given at *Semantics and Philosophy in Europe (SPE 7)*.

Esipova, Maria. 2019a. *Composition and projection in speech and gesture*: New York University dissertation. URL https://ling.auf.net/lingbuzz/004676.

Esipova, Maria. 2019b. Towards a uniform super-linguistic theory of projection. In Julian J. Schlöder, Dean McHugh & Floris Roelofsen (eds.), *Proceedings of the 22nd Amsterdam Colloquium*, 553–562. URL https://ling.auf.net/lingbuzz/004905.

Esipova, Maria. 2021a. Composure and composition. Ms., University of Oslo, submitted. URL https://ling.auf.net/lingbuzz/005003.

Esipova, Maria. 2021b. On not-at-issueness in pictures. *Glossa: A Journal of General Linguistics* 6(1). 83. doi: 10.5334/gjgl.1314.

Fruehwald, Joseph. 2016. Prolegomena to a grammar of knitting. URL https://jofrhwld.github.io/papers/knitting.pdf.

Fujita, Koji. 2017. On the parallel evolution of syntax and lexicon: A Merge-only view. *Journal of Neurolinguistics* 43(B). 178–192. doi: 10.1016/j.jneuroling.2016.05.001.

Goldin-Meadow, Susan & Diane Brentari. 2017. Gesture, sign, and language: The coming of age of sign language and gesture studies. *The behavioral and brain sciences* 40. e46. doi: 10.1017/S0140525X15001247.

Greenberg, Gabriel. 2018. Content and target in pictorial representation. *Ergo* 5(33). 865–898. doi: 10.3998/ergo.12405314.0005.033.

Greenberg, Gabriel. 2019. Tagging: Semantics at the iconic/symbolic interface. In Julian J. Schlöder, Dean McHugh & Floris Roelofsen (eds.), *Proceedings of the 22nd Amsterdam Colloquium*, 11–20.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & S. Jay Keyser (eds.), *The view from Building 20*, 111–176. Cambridge, MA: MIT Press.

Hess, Sally & Donna Jo Napoli. 2008. Energy and symmetry in language and yoga. *Leonardo*

41(4). 332–338. doi: 10.1162/leon.2008.41.4.332.

Hockett, Charles F. 1959. Animal "languages" and human language. *Human Biology* 31(1). 32–39. URL https://www.jstor.org/stable/41449227.

Hunter, Julie. 2019. Relating gesture to speech: reflections on the role of conditional presuppositions. *Linguistics and Philosophy* 42(4). 317–332. doi: 10.1007/s10988-018-9244-0.

Inkelas, Sharon. 2014. Non-concatenative derivation: Reduplication. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of derivational morphology*, 315–332. Oxford University Press. doi: 10.1093/oxfordhb/9780199641642.013.0011.

Kaplan, Abby. 2017. Exemplar-based models in linguistics. In *Oxford bibliographies in linguistics*, doi: 10.1093/OBO/9780199772810-0201.

Katz, Jonah & David Pesetsky. 2011. The identity thesis for language and music. Ms., Institut Jean-Nicod & MIT. URL http://ling.auf.net/lingbuzz/000959.

Khan, Sameer Ud Dowla. 2014. The intonational phonology of Bangladeshi Standard Bengali. In Sun-Ah Jun (ed.), *Prosodic typology II: The phonology of intonation and phrasing*, 81–117. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199567300.003.0004.

Lascarides, Alex & Matthew Stone. 2009. A formal semantic analysis of gesture. *Journal of Semantics* 26(4). 393–449. doi: 10.1093/jos/ffp004.

Lerdahl, Fred & Ray Jackendoff. 1983. *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Liddell, Scott K & Robert E Johnson. 1989. American sign language: The phonological base. *Sign language studies* 64(1). 195–277. doi: 10.1353/sls.1989.0027.

Maier, Emar. 2019. Picturing words: the semantics of speech balloons. In Julian J. Schlöder, Dean McHugh & Floris Roelofsen (eds.), *Proceedings of the 22nd Amsterdam Colloquium*, 584–592.

Maier, Emar & Sofia Bimpikou. 2019. Shifting perspectives in pictorial narratives. In M. Teresa Espinal, Elena Castroviejo, Manuel Leonetti, Louise McNally & Cristina Real-Puigdollers (eds.), *Proceedings of Sinn und Bedeutung 23*, 91–105.

Napoli, Donna Jo & Stephanie Liapis. 2019. Effort reduction in articulation in sign languages and dance. *Journal of Cultural Cognitive Science* 3(1). 31–61. doi: 10.1007/s41809-019-00027-3.

Patel-Grosz, Pritty, Patrick Georg Grosz, Tejaswinee Kelkar & Alexander Refsum Jensenius. 2018. Coreference and disjoint reference in the semantics of narrative dance. In Uli Sauerland & Stephanie Solt (eds.), *Proceedings of Sinn und Bedeutung 22*, 199–216. Berlin: Leibniz-Centre General Linguistics.

Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive science* 33(6). 999–1035. doi: 10.1111/j.1551-6709.2009.01047.x.

Pavese, Carlotta. 2019. The psychological reality of practical representation. *Philosophical Psychology* 32(5). 784–821. doi: 10.1080/09515089.2019.1612214.

Pfau, Roland, Markus Steinbach & Bencie Woll. 2012. *Sign language: An international handbook*. Boston: Walter de Gruyter. doi: 10.1515/9783110261325.

Pietroski, Paul M. 2018. *Conjoining meanings: Semantics without truth values*. Oxford University Press. doi: 10.1093/oso/9780198812722.001.0001.

Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199273829.001.0001.

Potts, Christopher. 2007. The expressive dimension. *Theoretical linguistics* 33(2). 165–198.

doi: 10.1515/TL.2007.011.

Rooth, Mats & Dorit Abusch. 2019. Indexing across media. In Julian J. Schlöder, Dean McHugh & Floris Roelofsen (eds.), *Proceedings of the 22nd Amsterdam Colloquium*, 612–624.

Sandler, Wendy & Diane Lillo-Martin. 2006. *Sign language and linguistic universals.* Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9781139163910.

Schembri, Adam C. 2001. *Issues in the analysis of polycomponential verbs in australian sign language (auslan)*: University of Sydney dissertation. URL http://hdl.handle.net/2123/6272.

Schlenker, Philippe. 2007. Expressive presuppositions. *Theoretical Linguistics* 33(2). 237–245. doi: 10.1515/TL.2007.017.

Schlenker, Philippe. 2018a. Gesture projection and cosuppositions. *Linguistics and Philosophy* 41(3). 295–365. doi: 10.1007/s10988-017-9225-8.

Schlenker, Philippe. 2018b. Iconic pragmatics. *Natural Language & Linguistic Theory* 36(3). 877–936. doi: 10.1007/s11049-017-9392-x.

Schlenker, Philippe. 2019. Prolegomena to music semantics. *Review of Philosophy and Psychology* 10(1). 35–111. doi: 10.1007/s13164-018-0384-5.

Schlenker, Philippe, Emmanuel Chemla & Klaus Zuberbühler. 2017. Semantics and pragmatics of monkey communication. In Mark Aronoff (ed.), *Oxford research encyclopedia of linguistics*, doi: 10.1093/acrefore/9780199384655.013.220.

Schlenker, Philippe & Jonathan Lamberton. 2019. Iconic plurality. *Linguistics and Philosophy* 42(1). 45–108. doi: 10.1007/s10988-018-9236-0.

Starke, Michal. 2010. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36(1). 1–6. doi: 10.7557/12.213.

Sterelny, Kim. 2012. Language, gesture, skill: the co-evolutionary foundations of language. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367. 2141–2151. doi: 10.1098/rstb.2012.0116.

Thompson, Arthur Lewis & Youngah Do. 2019. Defining iconicity: An articulation-based methodology for explaining the phonological structure of ideophones. *Glossa: a journal of general linguistics* 4(1). 72. doi: 10.5334/gjgl.872.

Weisberg, Jill, Shannon Casey, Zed Sevcikova Sehyr & Karen Emmorey. 2020. Second language acquisition of american sign language influences co-speech gesture production. *Bilingualism: Language and Cognition* 23(3). 473–482. doi: 10.1017/S1366728919000208.