Toward a Functional Grammar of Physics Equations

Kirk M. Williams, David T. Brookes
Department of Physics; California State University, Chico; Chico CA 95929

An area of student difficulty in introductory physics courses is how they use and reason with equations. We propose that part of this difficulty is due to meaning that is embedded in the structure of equations. As equations are manipulated, their structure and concomitant meanings change. As mathematics is considered the "language of physics," our starting point will be to propose that it has a grammar. As equations change form and meaning, they are doing so within a certain grammatical system. We will show how physics equations can be categorized and mapped to ideational clause types as devised by Halliday (1994). This mapping could be useful in relating the mathematical "language" used in physics to "natural language," benefitting physics instructors who are trying to understand the struggles of their students, and helping students to understand the rich meanings embedded in physics equations.

Keywords: physics, mathematics, linguistics, ontology, grammar

It is often stated, to the point of cliché, that "math is the language of physics." This is intuitive, and readily accepted by most. The concepts of physics can be explained without any mathematics whatsoever, but this approach results "...in an understanding of physics that is fundamentally different from physics as understood by physicists" (Sherin, 2001, p. 524). Certainly, for most who are looking for any kind of practical aptitude in these concepts, it is essential to be able to work with equations.

If, indeed, mathematics is the language of physics, what kind of a language is it? What is its system of grammar? Knowing this could be useful, especially for educators whose competence in this language has surpassed the need to think of its underlying structure. At such a high level of expertise, it can be difficult to truly understand what is causing students difficulties as they learn how to communicate and do physics with mathematics. If instructors could see how conceptually complex it really is to know what equations *mean*, perhaps they could better understand the struggles of their students and be better equipped to help. Research has been conducted into the lexical meaning of symbols in physics and how those diverse meanings pose both interpretative and epistemic difficulties for students (Torigoe & Gladding, 2011; Redish & Kuo, 2015). Others have examined the structure of equations themselves and how that structure facilitates or constrains physical reasoning (Sherin, 2001; Landy, Brookes, & Smout, 2014). Weinberg, Dresen and Slater (2016) have examined mathematics as a semiotic system used productively by students for meaning-making. But to the best of our knowledge, no real attempt has been made to develop a grammar of physics equations. Our goal in this paper is to lay the groundwork for this process. We are going to suggest that equations have fundamental spatial structure, ordering, and function that encodes underlying meaning and it is in this area that additional challenge arises for students. It should be noted that our focus is on physics equations, in particular those seen by students in lower-division undergraduate physics courses. The broader discussion about the grammar of mathematics as a whole is beyond the scope of our work.

This project began with the observation that common physics equations can be separated into different categories based on their meanings. As equations are rearranged or manipulated, these meanings change. For instance, $a=F/m$ (Newton's second law) is what we would call a *causal* equation; it has an effect - acceleration - on the left side, and a cause - force - on the right. Mass is an inhibitive contribution to the cause, as it is inversely proportional to the effect. Research has already empirically shown that equation users are sensitive to the ordering of causal equations (result left of the equals sign, cause on the right), and reversing the order is confusing or changes the meaning of the equation (Mochon & Sloman 2004). We claim this is prima facie evidence of equations having a grammatical structure.

Is Mathematics a Language?

Before proceeding, we must ask a simple question that - it appears to us - has no simple answer: Is mathematics legitimately a *language*? To address this, we will start in the broader realm of semiotics. Mathematics is most certainly a kind of semiotic system; it is vehicle for making meaning and communicating. Semiotics, to put it simply, is the study of signs. There are two predominant models of signs: the dyadic and the triadic. A dyadic sign would consist of a "signifier" and a "signified." The signifier could also be called the "sign vehicle" and the signified the "referent" (Noth, 1990). Essentially, there is a thing being represented and a way of representing it. The chief limitation of the dyadic model is that it lacks context. There are semantics and syntax, but no pragmatics (Ongstad, 2006). A triadic model, as devised by C.S. Peirce, would add what he called the *interpretant* to the previously described schema. A more commonly used term for this is *sense*, meaning there is someone or something "receiving" the sign and interpreting it. The triad is thus *sign vehicle*, *sense*, and *referent* (Noth, 1990).

Both of these models have played out in existing analysis of the semiotics of mathematics. A kind of dyadic model is proposed by Rotman (1988), in which the mathematical sign has the components *thought* and *scribble*. The two cannot be separated and be considered a true mathematical sign. In Rotman's mathematical semiosis, a person in essence *creates* a Mathematician, who then creates an Agent. Each of these take on the firstness, secondness, and thirdness, as devised by Peirce, as they proceed through the creation of a proof (Rotman, 1988). Ongstad advocates for a triadic model in which a sender and receiver are involved in the interpretation of content. The sign itself in this case is made up of the elements Symptom, Symbol, and Signal. These correspond to senders, "objects or states of affairs," and receivers, respectively (Ongstad, 2006).

The semiotics of mathematics is a rich topic. It is clear that in doing mathematics, we are engaging in some form of communication. But what of our treatment of mathematics as a language? Leibniz attempted to develop a universal language involving a "...*calculus ratiocinator*, a system of rules for the combination of semantic primitives" (Noth, 1990, p. 274). Frege's mathematical symbolism "embodies fundamental principles of reasoning based on an analysis of language" (Bouissac, 1998, p. 249). Rotman describes mathematical texts as being a "...mixture of natural and artificial signs...conventionally punctuated and divided up into what appear to be complete grammatical sentences…" (Rotman, 1988, p. 7). Ongstad gets more specific in proposing that mathematics could be "...a set of interrelated, semiotically different languages or sign systems" (Ongstad, 2006, p. 248). If one takes this perspective, it is reasonable to suggest that physics equations might be a language in their own right, semiotically different than the languages of pure mathematics or statistics (Redish & Kuo, 2015).

What is it that makes physics equations unique? At least part of it is that they do not consist of pure, abstract mathematical objects; rather, they use these objects to describe patterns in nature. This distinction is what shall characterize the ontology - and, as we shall see, the grammar - of physics equations.

Systemic Functional Grammar

If we are to devise a grammatical system of physics equations, it must - if it is to be useful - be analogous to one that is familiar to us. We will use our native language of English, but we will attempt to minimize applications of grammatical concepts that are not also applicable to other languages. For this kind of universality, we look to Functional Grammar, as devised by Halliday. It is a systemic theory, a "...theory of meaning as choice, by which language, or any other semiotic system, is interpreted as networks of interlocking options…" and it is "...functional in the sense that it is designed to account for how the language is **used**" (Halliday, 1994, p. xiii). Our focus has been primarily on the experiential aspects of the grammar, which look mostly at the clause. The clause in this framework has three different metafunctions (sometimes called components); ideational ("clause as representation"), interpersonal ("clause as exchange"), and textual ("clause as message") (Halliday, 1994). The ideational metafunction is the most appropriate for application to our mathematical "clauses" of interest. Physics equations do, after all, *represent* - or model - objects, interactions, systems, and states (Etkina, Warren, & Gentile, 2006).

The ideational metafunction models the clause as a process, within which there is an internal process (typically a verb), a participant (typically a noun), and, optionally, a circumstance (Halliday, 1994). For instance, in the sentence, "The girl caught a fish from a lake," "The girl" and "a fish" are participants, "caught" is a process, and "from a lake" is a circumstance. These classifications are quite broad, so different process types are used depending on the function of the clause. This type of categorization is called a *transitivity system*. Structure is "explained in terms of meaning." The three primary process types within the ideational component are material, mental, and relational processes. These are processes of doing, sensing, and being, respectively (Halliday, 1994).

For example, a material process might be our previous example (omitting the circumstance for brevity), "The girl caught a fish." In this case, we call "the girl" the Actor and "a fish" the Goal. These describe processes of *doing*.

In the case of a mental process, it is necessary to have a personified participant - something that can "sense" something else. In the sentence, "He likes it," "he" is classified as a Senser and "it" is a Phenomenon. "Likes" is the process. We sometimes use this language in physics to personify things like electrons, which we might say "want to be in the ground state" (Brookes & Etkina, 2007).

Relational processes are the most varied and complex, as they describe relationships in which things are identified, symbolized, or otherwise related to other things. The two main types of relational processes are attributive and identifying. The participant types associated with these are Carrier/Attribute and Identified/Identifier (sometimes called Token/Value), respectively. The former treats a participant as a member of a category, while the latter *identifies* the participants as each other, and is thus reversible ("Alice is wise" vs. "Alice is the wise one") (Halliday, 1994). In addition, the relational process has three subcategories: Intensive ('x is a'), circumstantial ('x is at a'), and possessive ('x has a'). These can each be combined with either "attributive" or "identifying" to form such combinations as "circumstantial identifying" or "possessive attributive." Distinctions like this will be quite useful in formulating a kind of transitivity system for physics equations.

Finally, our brief summary of some important aspects of functional grammar must include a discussion of what goes on "below the clause." At this level, the ideational component splits into two categories: Experiential and logical. These turn out to be two different ways to examine phrases and groups within a clause, and the ordering of functional elements within a group. For instance, let's look at the experiential structure of the nominative group "those two old diesel trucks." It exhibits the typical ordering of elements: Deictic, Numerative, Epithet, Classifier, and Thing. To arrange the sentence in any other way would not make sense. If we look at the same group's logical structure, we would call "trucks" the Head and everything else the Modifier. Each word is then assigned a Greek letter, starting on the right with "trucks" and moving to the left. We would thus *read* this group's logical structure from left to right: Modifier (ε, δ, γ, β), Head (α). Conceptually, this ordering is characterized as moving "...from the kind of element that has the greatest specifying potential to that which has the least…" (Halliday, 1994, p. 187) This type of analysis could be effective in characterizing the order of elements in mathematical "groups" as well. There are clearly certain consistent tendencies, like putting numerals before constants, which are then put before variables. Our focus here is less on the group and more on the clause, as we aim to set up a transitivity structure for our equations. However, the ways in which mathematics is at least "like" a language continue to unfold; it does not appear to be a superficial connection.

Ontology, Grammar, and Interpretations of Equations

An important concept in this discussion is that of ontological "trees," as devised by Chi, Slotta, & de Leeuw (1994) and later modified by Brookes and Etkina (2007). Chi et al. proposed that people separate the world into three primary ontological categories (trees), each having its own subcategories (branches). These are Matter, Processes, and Mental States. When an idea or entity is initially conceived to belong to one of these categories, and then must be moved to another, this is called *conceptual change*. Topics that require this kind of shifting exhibit a kind of "incompatibility" of conception and tend to be more difficult to learn. This is part of what is called the Incompatibility Hypothesis. Many science concepts require the learner to continually alternate between categories, which creates exceptional difficulty (Chi et al., 1994).

The version of this model adapted specifically for the language of physics by Brookes and Etkina changes the category of Mental States to the more general States (Brookes & Etkina, 2007). Etkina, Warren, & Gentile devised a taxonomy of physical models, which comprised of models of objects, interactions, systems and processes (split into causal and state equations) (Etkina et al., 2006). These were mapped to the ontological categories of matter processes and physical states, in part to understand the prominent use of metaphorical language in how physicists talk about physical ideas (Brookes & Etkina, 2007). For example a physicist might say "Energy flowed into the system by heating." In this sentence "energy" is the *matter*, "flowed" is the *process*, and "by heating" is a circumstance that elaborates the nature of the *process*. On the other hand, a physicist could say "Heat flowed into the system." In this case, "heat" has shifted in its grammatical function (from circumstance to participant) and likewise has shifted its ontological category from elaborating the *process* to being categorized as *matter*.

Another important precedent for our work is the concept of *symbolic forms*. Symbolic forms are what Sherin (2001) describes as "knowledge elements" with two components: A symbol template and a conceptual schema. The symbol template component is primarily how Sherin distinguishes the forms. This is an abstraction of a mathematical expression in which symbols are replaced primarily with shapes ($\square$), generalized variables (x) and ellipses (...), so that the focus is on the structure of the expression, rather than its specific content. For example, the symbolic form "balancing" is represented with the symbol template $\square = \square$. "*Identity*" is represented with $x = […]$. Sherin presents a "semi-exhaustive list" of these forms, and suggests that its organization into "clusters" is "...primarily for rhetorical purposes - not to reflect any psychological grouping of the elements. However, within a given cluster, the various schemata tend to have entities of the same or similar ontological type" (Sherin, 2001, pp. 505-506). We suggest that the reason these clusters are not clearly defined is because of the level of abstraction used in the model of symbolic forms. The meanings of the mathematical structures cannot be adequately understood when their "participants" have been generalized and removed from their processes. Sherin seems to be primarily analyzing the constituent structure (Halliday, 1994) of the equations' *orthography*, not their grammar. That is, we see how this "language" is typically written down, but not how it is used. For example: $\vec{p} = m\vec{v}$ fits into the *identity* template and is recognized as such in physics, but so does $N = mg$, an equation that does not represent *identity*.

Among Sherin's references is Anna Sfard, who has written at length about the meaning of the equals sign. She has suggested that although "*...there is a deep ontological gap between operational and structural conceptions*… they are in fact *complementary*" (Sfard, 1991, p. 4). This duality of object vs. process is found in many forms in the mathematics education literature, but Sfard's comparison of this distinction to the "complementarity" of waves and particles in quantum mechanics is unique. This illustrates the subtlety and ambiguity of the equals sign, our grammatical "process."

Mapping Grammar to Equation Types

Taking into consideration all of the literature reviewed above, the most potent for analyzing the meanings of physics equations has been Halliday's functional grammar. It works surprisingly well to simply map ideational process types to equation categories. This mapping is certainly not one-to-one, but it offers useful distinctions and is remarkably consistent with the way these equations are used in (for example) Knight's popular undergraduate physics textbook (Knight, 2004).

We started the mapping by dividing equations in physics into three broad categories based on three distinct meanings of the equals sign. Building on prior work (Keiran, 1981; Sherin, 2001; Redish and Kuo 2015) we recognized that an equals sign in physics can mean "is," "is equal to," or "is a result/consequence of." The second key to mapping equations to grammar is to examine how the equation functions in relation to the words that surround it. In other words: equations in physics *cannot* be separated from the surrounding (English) sentence if we want to understand their full meaning. We will present our analysis based on a commonly used University-level, introductory physics textbook (Knight, 2004).

In our framework, equations in which the equals sign means "is" belong to a category we called "Operational Definition," similar to Sherin's *identity* template. For example, acceleration $\vec{a} = \frac{d\vec{v}}{dt}$, or momentum $\vec{p} = m\vec{v}$ operationally define useful physical quantities in terms of how they are measured. These correspond to grammatical relational intensive processes. We take this to generally also mean these processes are identifying rather than attributive, because mathematics does not have quite the same issues with reversibility and active vs. passive voice that we have in English. For example, when dealing with momentum, Knight writes, "The product of a particle's mass and velocity is called the momentum of the particle: momentum = $\vec{p} = m\vec{v}$" (Knight, 2004, p. 262). This is an intensive identifying clause; it is reversible, and it serves to assign an identity to the signs p and mv. The equation itself does essentially the same thing. The two signs are interchangeable - the equation serves to identify.

Next, we identified a category of equations where the equals sign reads "is equal to". Equations in the "Is Equal To" category map to relational circumstantial processes. Each of these is true only within certain specific circumstances. One example from Knight: "If the angle θ is such that $\Delta r = d\sin\theta = m\lambda$, where m is an integer, then the light wave arriving at the screen from one slit will be *exactly in phase* with the light waves arriving from the two slits next to it" (Knight, 2004, p. 938). In this sentence $d\sin\theta = m\lambda$ functions as a grammatical circumstance. On the other hand, that equation, presented on its own on (for example) an equation sheet, lacks any indication of its circumstantial nature. We hypothesize that an expert seeing this equation implicitly sees the surrounding context as well, even when it is absent. The sentence quoted doesn't come out to be a single circumstantial clause, but it doesn't need to. As we have noted, this mapping from the grammar of language to the grammar of physics equations is not one-to-one. The distinguishing characteristic of these equation types is their being tied to circumstance, sometimes in a subtle way. For example, $N = mg$ (another equation that falls into the "is equal to" category) is only applicable when an object of mass m is resting on a level/horizontal surface, close to the earth's surface and ignoring the fact that our rotating (Earth's) reference frame is slightly non-inertial. Halliday writes of circumstantial identifying processes in which the participants act as the circumstantial element: "The relation between the participants is simply one of sameness; these clauses are in that respect like intensives, the only difference being that here the two halves of the equation - the two 'participants' - are, so to speak, circumstantial elements in disguise" (Halliday, 1994, p. 131). This true of an equation like $\lambda = v/f$. In order for this to be true, there has to *be* a frequency f to create a wave with a wavelength λ constrained to be traveling at a velocity v through a medium.

The third meaning of the equals sign ("is a result/consequence of") defines a category of equations we have called "Causal." These equations represent material processes - processes of *doing*. Much as we have an Actor and a Goal in such grammatical processes, causal equations have what we shall call a Result (or Change) on one side of the equals sign, with an Agent and, optionally, an Inhibitor on the other. In the ubiquitous equation that represents Newton's second law, $\vec{a} = \frac{\sum \vec{F}}{m}$, "$a$" is our Result (*defined* to be $\frac{d\vec{v}}{dt}$: a rate of change in velocity with respect to time), while "$\sum \vec{F}$" is the Agent and "m" is the Inhibitor. These kinds of relationships are some of the clearest, as we can see in Knight's words: "A force applied to an object causes the object to accelerate" (Knight, 2004, p. 120). It is interesting to note that some controversy about the causal form of Newton's second law exists. Although many textbook authors readily acknowledge forces exerted on an object cause the object to accelerate, Newton's second law is frequently written in a form that contradicts that causality: $\sum \vec{F} = m\vec{a}$. Future work needs to examine whether there is indeed a student difficulty that arises from the apparently inconsistent ways in which Newton's second law is presented and understood by experts. Another example of a causal equation with the Inhibitor absent would be the first law of thermodynamics: $\Delta U = Q + W$. Here Q and W represent the two possible agents (heating and work) that result in the change in internal energy ΔU of the thermodynamic system.

Implications and Future Work

Students in beginning physics courses face many challenges, but perhaps the most daunting for them is the use of equations. A substantial contributing factor to this area of student difficulty is a sense of what equations *mean*. This problem is both lexical *and* ontological in nature, and if we are to understand the ontological challenges that students face, we need to understand the grammar of physics equations. We suggest that in order to understand equations in action we need to understand how their meaning shifts as they are rewritten and manipulated. For example, a student might start out with a causal form of Newton's second law: $\vec{a} = \frac{\sum \vec{F}}{m}$ and at some later point rewrite the same equation as $F = ma$ to find the value of a particular force in order to solve a specific physics problem. In manipulating the equation this way, force has shifted grammatically (and consequently, ontologically) from being an Agent to an entity that can be determined from other physical quantities. Grammatically the equation has shifted from the material process category to being a circumstantial clause. We hypothesize that, in a way that is analogous to spoken or written language, the meaning and function of entities in an equation are constantly shifting as the equation is re-organized, and manipulated. A key part of reasoning with mathematics (Redish and Kuo, 2015), is being comfortable with these shifts, just as a native speaker of a language is comfortable turning verbs into nouns and noun into verbs in a way that is communally understood by other native speakers of that language. In short, the mark of native speakers of a language is their ability to play "fast and loose" with the lexico-grammatical interaction of that language and still engage in meaningful communication. A detailed exposition of this should be followed up upon in future work.

Educators often take the reasons for students' difficulty with equations for granted because of their experience and expertise. If educators could see how complex the meanings of common physics equations are, they could perhaps be more equipped to help students make sense of them. If we are willing to look at these equations as a part of the "language of physics," as most already do, we can treat them as clauses in this language. Halliday's Functional Grammar is a useful tool in making meaningful comparisons between different types of equations and types of clauses in English. Mapping between our proposed categories of physics equations and Halliday's transitivity system for ideational clauses works exceptionally well as a theoretical framework to understand meaning-making with equations

The theory as presented is a sketch. It has the potential to be fine-tuned with more analysis of physics textbooks, as well as deeper research in linguistics and perhaps other fields. Involving experts in other fields, such as linguists, educators, psychologists, mathematicians, and more, could be of immense benefit to the theoretical framework.

Finally, it will be necessary to devise experiments from which we can extract data to help determine the useful applications of this idea. Pedagogical strategies involving the theory must be developed and then tested, perhaps by surveying large groups of students and/or examining smaller groups working and reasoning with equations. The implementation of this theory into curriculum could be subtle, where the instructor could simply repeatedly emphasize the meanings of equations, or more overt, where the students are explicitly made aware of the equations' grammatical structure and the implicit meaning associated with that structure.

References

Bouissac, P. (Ed.). (1990). *Encyclopedia of Semiotics*. USA: Oxford University Press.

Brookes, D. T., & Etkina, E. (2007). Using conceptual metaphor and functional grammar to explore how language used in physics affects student learning. *Physical Review Special Topics Physics Education Research*, *3*(1), 010105.

Chi, M. T. H., Slotta, J. D., & de Leeuw, N. (1994). From Things to Processes: A theory of conceptual change for learning science concepts. *Learning and Instruction*, *4*(1), 27–43.

Etkina, E., Warren, A. A., & Gentile, M. (2006). The Role of Models in Physics Instruction. *The Physics Teacher*, *44*(1), 34 – 39.

Halliday, M. A. K. (1994). *An Introduction to Functional Grammar* (2nd ed.). London: Edward Arnold.

Knight, R. D. (2004). *Physics for Scientists and Engineers: A Strategic Approach* (4th ed.). San Francisco: Pearson Addison Wesley.

Landy, D., Brookes, D., & Smout, R. (2014). Abstract numeric relations and the visual structure of algebra. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1404–1418.

Mochon, D., & Sloman, S. A. (2004). Causal models frame interpretation of mathematical equations. *Psychonomic Bulletin & Review*, *11*(6), 1099–1104.

Noth, W. (1990). *Handbook of Semiotics*. Bloomington & Indianapolis: University of Indiana Press.

Ongstad, S. (2006). Mathematics and Mathematics Education as Triadic Communication? A Semiotic Framework Exemplified. *Educational Studies in Mathematics*, *61*(1/2), 247–277.

Redish, E. F., & Kuo, E. (2015). Language of Physics, Language of Math: Disciplinary Culture and Dynamic Epistemology. *Science & Education*, *24*(5-6), 561–590.

Rotman, B. (1988). Toward a semiotics of mathematics. *Semiotica*, *72*(1/2), 1–35.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, *22*(1), 1–36.

Sherin, B. L. (2001). How Students Understand Physics Equations. *Cognition and Instruction*, *19*(4), 479 – 541.

Torigoe, E. T., & Gladding, G. E. (2011). Connecting symbolic difficulties with failure in physics. *American Journal of Physics*, *79*(1), 133–140.

Weinberg, A., Dresen, J., & Slater, T. (2016). Students' understanding of algebraic notation: A semiotic systems perspective. *Journal of Mathematical Behavior*, *43*, 70–88.