On *morpho*-syntax

Daniel Siddiqi
Carleton University

0. Introduction

While overshadowed a bit by the advent of the Minimalist Program, Bare Phrase Structure, and Merge (MPP; Chomsky 1995) the rise of a realizational morphology-syntax interface (Anderson 1992, Beard 1995, Halle & Marantz 1993, 1994, Stump 2001, etc) during that same period (the early 1990s) has had similarly profound effects on the enterprise of generative grammar. Of course, realizational morphology-syntax interfaces are in no way limited to the Chomskyan generative tradition. For example, a fruitful contemporary research program maps Paradigm Function Morphology (PFM; Stump 2001, 2016) to Lexical Functional Grammar (LFG; Bresnan & Kaplan 1984). However, in the context of this workshop, the main relevant expression of this in MPP is Distributed Morphology (DM, Halle & Marantz 1993, 1994), which has risen over the last quarter century to be one of the dominant Minimalist frameworks. Its addition to MPP has paid huge dividends to the model, not the least of which is opening access to a large swath of North American languages as DM has equipped MPP with ample tools required for investigation into highly synthetic languages.

An important aspect of the rise of DM is the blurring of lines in generative grammar. Within the Chomskyan tradition, the division of morphology and syntax has largely been erased, meaning Chomskyan syntacticians are often morphologists and DM-morphologists are often syntacticians. This is a massive growth in the scope of the enterprise (or more precisely a return to a previously large scope that had been narrowed circa Chomsky 1970). For example: MPP, by way of DM, is now crucially concerned with the nature of the morphology-phonology interface. Stem allomorphy, productivity, and blocking, which were once the exclusive domain of morphology, are now concerns for syntax. Whole word storage, word processing, and frequency effects are now within MPP's domain. Templatic morphology, defective paradigms, and cross-class syncretisms are things MPP ought to have explanations for. Similarly, DM has become a model of lexical semantics, especially recently. And again, MPP now inherits the onus to provide explanations to complement the decades of lexical semantics research, such as argument structure, polysemy, and event structure.

In this type of venue, it is tempting to bemoan the specialization and lack of awareness of the literature of researchers in this type of sprawling interface research program. Certainly, I will do some of that here, but my main contribution to this workshop is a discussion of crucial ways the move from "syntactic" theory to "morphosyntactic" theory has changed the mission of generative grammar and to what extent practitioners have kept pace. I focus here on a brief survey of five things that seem especially pressing to me.

1. On exceptions and generalizations

One of the key markers of morphology is that it is the main domain for exceptional behaviour. For example, the generalization about morphology is that it is largely concatenative, showing headed hierarchal structure similar to syntax. However, stem allomorphy and other non-affixal morphological processes that modify the base make up a significant portion of studied

phenomena, especially among more frequent forms. Similarly, most (affixal) morphology is semantically transparent. But, the complex word is famously the domain for most non-compositional meaning (often unironically called lexicalization). Even in the domain of phonology, morphophonemic alternations such as tri-syllabic shortening or velar softening stand apart from the otherwise fully regular and fully productive phonology.

The relative size of the class of exceptional phenomena within the domain of morphology is what makes it significant. Non-concatenative morphology comprises such a large class of data that it is not an exaggeration to claim that the field of morphology is effectively divided over the metatheoretic concerns regarding dealing with it. Item-and-arrangement models of morphology seek to treat non-affixal morphology as truly exceptional and limit their scope to primarily account for concatenative morphology. This limits the formal mechanism they require to a simple concatenative mechanism not unlike Merge, thus increasing restriction and decreasing power). The flip side of this, of course, is that all morphology has to be treated as affixal (thus the need for transfixes and superfixes and the like) or it has to be treated as exceptional (and essentially listed). This trade-off reduces the empirical coverage of item-and-arrangement models. However, one significant gain morpheme-based models is a claim to parsimony. In this view, syntax and morphology can both be reduced to headed hierarchal structures.

Item-and-process and word-and-paradigm models of morphology treat the preponderance of concatenative morphology as epiphenomenal and seek to treat non-affixal morphology as the generalization than needs to be accounted for. In these models, morphological rules or paradigm functions are powerful enough to generate non-concatenative morphology. Thus, the power of these models is generally greater than needed to account for the attested patterns of morphology in the world and is certainly greater than that needed to account for concatenative morphology. Indeed, in these models, what appears to be an affix is almost never treated as one. Rather, it is a stem change that happens to be at the periphery of the word form. In these cases, the fact that most morphology appears to be concatenative is treated as an artifact of history (see Stump 2001 for discussion).

In essence, item-and-arrangement models sacrifice empirical coverage by treating non-affixal morphology as exceptional, but gain restriction and the corresponding falsifiability. The other two gain empirical coverage but at the price of increased power.

Of course, this discussion should be very familiar. Chomsky (1970) in no small part argued to eliminate these types of exceptional data from the mission of syntactic generative grammar. Lieber (1992) made similar arguments to eliminate them from a model of the morphosyntax as well. Since Chomsky (1995), MPP has put a primary focus on the metatheoretic principles of restriction and elegance. As it has done so, it has systematically sacrificed empirical coverage to limit its power, eventually discarding adjuncts and head movement, for example. The restriction it has gained is frequently touted by its practitioners as its main appeal, especially as its competitors such as HPSG, LFG, and CxG have sacrificed restriction for empirical coverage and computational/cognitive power.

So it is a remarkable that DM undoes a lot of those moves.

DM spends an extraordinary amount of research on exceptions and creates increasing numbers of mechanisms to account for this exceptional behaviour. It is not without significant irony that I

point out that root suppletion and stem allomorphy is one such topic where practitioners of DM invest significant effort (see for example Siddiqi 2009, Haugen & Siddiqi 2013, 2016, Harley 2014, Harley & Tubino-Blanco 2014 among many many others) and create powerful mechanisms (such as default fusion or readjustment rules) to account for phenomena that are clearly exceptional and would be removed from the enterprise of generative grammar by the standards of such as Chomsky (1970) and Lieber (1992). Similarly, recent accounts for level ordering data (see for example Newell 2016; cf. Kiparsky 1982) again capture data with very low productivity (potentially none) yet suggest significant changes to the power of the productive grammar.

This monster of power creep is two-headed. Besides accounting for phenomena that is clearly exceptional, DM eagerly accepts mechanisms that were cast aside from the pre-spell-out branch of syntactic theory, such as head movement and lowering while also adding powerful transformational mechanisms such as rebracketing (see for example Radkevich 2014) and local dislocation (see for example Embick 2007).

The sheer number of these exception-capturing mechanisms would drastically reduce restriction and increase the power of the model by itself, but this is further magnified by the fact that these mechanisms themselves are also frequently powerful. Impoverishment and readjustment have existed for as long as the theory has. Recently, allosemy, which is essentially readjustment except for lexical semantics instead of phonology, has been increasingly en vogue (see for example Harley 2014, Marantz 2014). Readjustment and allosemy are especially powerful, so I will give them a section of their own below.

MPP (and the Chomskyan tradition more generally since at least 1970) has taken as its driving principle that restriction, elegance, parsimony, and other such concerns are prioritized over empirical coverage, especially when it comes to phenomena that are clearly exceptional. MPP stands nearly alone among formal models of syntax in prioritizing these concerns. DM inherits this mission from MPP, but while MPP continues to emphasize this restriction, it seems like DM when taken as a whole, loses sight of this mission. To some extent, this is just the natural effect of power creep. Each theorist who proposes a small increase in power has not compromised the whole system, but when all of us are doing this, we quickly arrive where we are now: with a PF path that is congested with powerful devices that threaten the claim to restriction.

But the other reason that this increase of complexity is happening is the types of questions we are asking: specifically about morphology. The field of generative morphology has known since its inception that its generalizations license a simple, elegant model of concatenative morphology and syntax—a model that DM aims to be. But its exceptions are a large enough class that they license a more robust word-based approach, such as the approach taken by PFM. The word-based hypothesis (Aronoff 1976) treats explanations for the exceptional such as stem allomorphy as the measure as the minimum necessary power. They are very good at this. DM needs to remember that its appeal, and the appeal of any morpheme approach, is its limited power. This entails treating exceptional morphology as exceptional. Otherwise, it will just be worse at doing what word-based models are already doing but without any claims to restriction and elegance.

2. On "conceptual" arguments

What I hope I did in Section 1 was make a nuanced, metatheoretical argument couched in a solid history of philosophy of science. I.e. a "conceptual argument". Generative grammar, like all

sciences, relies on metatheory as the backbone of the entire enterprise. Indeed, MPP is essentially a mandate to put metatheoretic concerns at the forefront of linguistic theory. It is odd then that even a cursory glance at the literature in DM will reveal "setting aside conceptual considerations" as a refrain. To be sure, given the relatively very small amount of data we have, there is definitely a place for arguments that the field of linguistics should be focusing more on developing models with sweeping empirical coverage than developing models that forefront metatheoretical concerns. Kaplan (1987) very convincingly made this argument in defense of very complicated and powerful modular models such as LFG. A similar argument is made by Haspelmath (2013) in favor of non-generative, "nonaprioiristic" approaches to comparative syntax. These types of arguments make sense when made to argue on behalf of a model or framework or program that has aimed at putting empirical coverage ahead of metatheoretic concerns. These arguments are somewhat counterintuitive when used within DM since DM appears at the nexus of a history of morpheme approaches to morphology (which by their very nature prioritize parsimony and restriction over empirical coverage) and Minimalist syntax (which explicitly prioritizes parsimony and restriction). Metatheoretic concerns are fundamental to DM. They are in its very blood.

The most compelling case against conceptual arguments that I have seen within DM comes from Embick (2014), where he argues against conceptual arguments being used to favor "insertion into non-terminals" models over "morpheme-insertion only" models, so I quote a bit of it here:

> Since it appears to allow for a grammar with fewer mechanisms, INT might have a conceptual advantage over MIO; at least, to the extent that this kind of accounting is taken at face value as a valid assessment of parsimony…. In any case, conceptual arguments about which of INT and MIO has more or less machinery provide guidelines for research, but are not decisive, and must play a secondary [role] to questions about where the two theories differ empirically. …MIO is superior to INT on empirical grounds, in a way that trumps (potential) conceptual concerns. (Embick 2014)

In the case of Embick (2014), I happen to have disagreed with the relevant empirical arguments (Haugen & Siddiqi 2016), but the argument for prioritizing empirical concerns is a compelling one…except that if you follow that to its logical conclusion, it means adopting a word model of morphology that shares many of the underlying assumptions of DM but has much greater empirical coverage (and the corresponding power) (one obvious candidate I mentioned before is PFM). DM and MPP have explicitly stated, repeatedly, by their very design, that conceptual concerns must play a primary role, not a secondary one.

An ancillary concern I have here is the role a potential dismissal of "conceptual considerations" can have on the review process. Having a propensity for the conceptual argument myself and having been an editor on several occasions, I have seen more than my share of reviews that actively diminished the value of metatheoretic concerns. While it is certainly true that conceptual arguments should be disregarded in the face of overwhelming counter-evidence (even the simplest model has no value if it doesn't describe the world), conceptual arguments have a significant importance to MPP and DM. We risk losing our identity if we are too quick to disregard them.

3. On readjustment and allosemy

I have argued against readjustment rules for most of my (very short) professional career. But these arguments have usually come from the point of view of a morphologist and a syntactician. It should be lost on nobody though, that nearly every contemporary incarnation of readjustment rules assumes they are *phonological*—that they properly belong to the phonological component of the grammar. Similarly, the very recent rise of allosemic alternations, as conceptualized by Harley (2014) or Marantz (2014) are intended to capture polysemy, which has always been one of the most researched areas of *lexical semantics*.

Both readjustment and allosemy have in common that they are aggressively powerful. Both have the power to map one root to vastly different LF and PF realizations given a particular environment. As Haugen & Siddiqi (2013, 2016) point out, readjustment is assumed to account for alternations such as *think-thought, seek-sought,* and *bring-brought*, which involve a phonological rule powerful enough to take /ɪŋk/, /ɪk/, or /ɪŋ/ as its input and then output /ɔ/. This is an extraordinarily powerful phonological rule: the type of which would never appear in phonological literature. Similarly, the current way we approach allosemy, as seen in Harley (2014), employs radically different and independent mappings from a root to its semantics. For example, Harley (2014) proposes that the root realized by *–ceive*/*-cept-* maps to "think" in the context of *con-* while mapping to "fake" in the context of *de-*. Marantz (2014) and others extend this type of licensing to functional items as well.

Put succinctly, these are examples of morpho-syntacticians imposing brute-force, radical replacement operations on *a different component of grammar*—rules of a variety that would never be proposed by practitioners in those particular components. It's not surprising that there has been pushback about this: Bermudez-Otero (2013) is in part a scathing rebuke of readjustment, and Ramchand (2015) forcefully rejects allosemy from the view of a constructivist (a position on lexical semantics that DM assumes). Ramchand's (2015) arguments are as compelling as they are obvious: the gains seem minimal (syntactic independence from lexical semantics) while the metatheoretic costs are enormous (adding another listing frame to the grammar, destroying learnability through syntactic-semantic bootstrapping, and obliterating any and all generalizations about the nature of polysemy).

Why are we doing this with the syntax? Is it not weird that we are ignoring decades of generalizations and best practices in these other components so that we can account for their data in ways they never would (via brute force and stipulation) and in ways we never would for own data?

4. On words and morphemes

I promised to "bemoan the specialization and lack of awareness of the literature of researchers in this type of sprawling interface research program". I'll do that here. When grad students ask me for reading on how to become a morphologist in DM, I hand them three books: Aronoff (1976), DiSciullo & Williams (1987), and Stump (2001). This is because, in my experience, grad students entering into the enterprise of being a DM morphologist have almost no awareness of the morphological literature outside of DM. Hearing that the existence of morphemes is a hotly debated topic *in the field of morphology* seems to them almost analogous to hearing tales of the boogeyman. This is not the fault of grad students or their supervisors. This is the fault of the literature in DM.

As we mentioned in our editors' note to Siddiqi & Harley (2016), the morphological literature in DM is strikingly, almost mind-boggling, insular. There are some standout counter-examples (see for example many papers and books from David Embick), but those usually argue that DM is globally preferable to word-based approaches. Seldom does the DM literature draw on insights from the word-based literature about particular phenomena. This tendency might be true of every model of every field of linguistics. It is certainly true of MPP more generally. Though, in MPP, there is at least a claim to predominance in syntactic research. How you go about delineating DM syntacticians from DM morphologists is outside of my ability set and thus so is counting morphologists and guessing at percentages, but it doesn't strike me as true that DM has the same claim to predominance in the field of morphology, especially given the immense literature in word-based theory. This is to say nothing of the lion's share morphology and word processing have in the psycholinguistics literature.

My discussion above about metatheoretic concerns of power and restriction do not occur in a vacuum. When DM does explicitly engage the greater literature on morphological theory, it is to make strong claims about restriction, power, economy, and parsimony. It inherits most of these claims by way of the morpheme hypothesis. It also inherits its weaknesses (stem allomorphy, defective paradigms, cran-morphs, bound stems, etc). DM practitioners should always be aware of this. Increasing the power of DM jeopardizes the strength of morpheme models. Eventually, the increased power will undermine the restriction of DM, and restriction will no longer differentiate it from paradigmatic realizational models such as Anderson (1992) and Stump (2001).

This all stops well short of engaging the most important part about ignoring such a large cross-section of the literature: there are very good linguists with very good insights about the same phenomena that we work on. We shouldn't be ignoring their insight any more than they should ignore ours because we are also syntacticians.

5. On psycholinguistic evidence

Almost as a throw away remark above, I commented that the study of word-formation makes up a disproportionate share of psycholinguistic research. In DM, we almost never make reference to this research (though Marantz 2005 and Pfau 2009 are solid counter-examples). There are good reasons for this, of course. Foremost of these is the claim that DM is model of competence. It's not super easy to see what studies of word processing effects or frequency or productivity have on a competence model of morphology, but it is certainly not nothing. For example, we regularly make aggressive claims of morphological decomposition. It seems like psycholinguistic evidence can inform these claims. DM is indeed couched in a literature of syntactic competence and in that domain it is much easier to conclude that psycholinguistic evidence doesn't inform the theory to a great degree. But a competence model is meant to be a model of linguistic knowledge and psycholiguistic evidence certainly tells us about speaker knowledge.

It is certainly very weird that DM shows less reliance on experimental evidence than the rest of morphology. Phonetics and phonology conferences are increasingly dominated by experimental evidence. Correspondingly, morpho-phonology increasingly relies on such evidence. Word-based models and adaptive discrimitive models (see for example Blevins et al 2016) are happy to incorporate psycholinguistic evidence. There are several claims within DM that can be tested experimentally. Decomposition and parsing are certainly the most prominent phenomenon

begging for experimental confirmation in a theory that increasingly proposes inflated numbers of morpheme boundaries and heavily articulating functional structure. Furthermore, it seems that DM predicts that increased processing time for forms that involve readjustment. This seems like a testable claim, though one that is certainly confounded by the fact that readjusted forms are usually high-frequency.

Psycholinguistic research methods are certainly not without their risks. For example, psycholinguistic evidence, especially in morphology, relies heavily on controlling for frequency effects. Frequency controls are dependent on the corpora the frequencies are drawn from, which have in turn made design choices and employed data collection techniques that have significant effects. The net result is that corpora can have butterfly-effect style consequences to morphological theory. See Swanson (2016 and following) for discussion.

6. Conclusions

More than once, I've heard the refrain at a morphology conference that "DMers are not morphologists but syntacticians." This is clearly not true. DMers are morphologists in every way that matters. But this claim is also not without merit. DM has a tendency to not engage the rest of the morphological literature. It also has a tendency deal with typical morphological concerns (i.e. exceptions) in ways that disregard chief morphololgical metatheoretical concerns. These are pretty significant objections. Since the conference is intended to be a "state of field" reflection on generative grammar (the Chomskyan enterprise), and that has returned to the pre-Remarks state of including morphology, the way Minimalism interfaces with the morphological literature and other morphological theories seem to be of chief concern and worth a moment's reflection.

References

Anderson, Stephen. 1992. A-morphous Morphology. Cambridge: Cambridge University Press

Aronoff, Mark. 1976. Word Formation in Generative Grammar. Cambridge: MIT Press.

Beard, Robert. 1995. Lexeme-Morpheme Base Morphology; a General Theory of Inflection and Word Formation. Albany, NY: SUNY Press.

Bermúdez-Otero, Ricardo. 2013. The Spanish lexicon stores stems with theme vowels, not roots with inflectional class features. Probus 25(1): 3-103.

Blevins , James, Farrell Ackerman, Robert Malouf, & Michael Ramscar. 2016. Morphology as an adaptive discrimitive system. In Daniel Siddiqi and Heidi Harley, eds., Morphological Metatheory. Amsterdam: John Benjamins

Chomsky, Noam. 1970. Remarks on Nominalization. In R. Jacobs & P. Rosenbaum (eds), Readings in English Transformational Grammar, 184–221. Waltham, MA: Ginn.

Chomsky, Noam. 1995. The Minimalist Program. Cambridge, MA: MIT Press.

Di Sciullo, Anna Maria & Edwin Williams. 1987. On the Definition of Word. MIT Press, Cambridge MA.

Embick, David. 2007. Linearization and Local Dislocation: Derivational mechanics and interactions. Linguistic Analysis 33:3-4, 303–336

Embick, David. 2014. On the targets of phonological realization. Proceedings for the MSPI Workshop at Stanford University.

Halle, Morris and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger, ed. Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge, MA: MIT Press.,

Halle, Morris, and Alec Marantz. 1994. Some key features of Distributed Morphology. In Andrew Carnie and Heidi Harley, eds., Papers on phonology and morphology, 275–288. Cambridge, MA: MIT Working Papers in Linguistics 21.

Harley, Heidi. 2014. On the identity of roots. Theoretical Linguistics 40: 225–275.

Harley, Heidi & Mercedes Tubino Blanco. 2013. Cycles, vocabulary items, and stem forms in Hiaki. In Distributed Morphology Today: Morphemes for Morris Halle, ed. by Ora Matushansky and Alex Marantz, pp. 117-134. Cambridge, MA: MIT Press.
Haspelmath, Martin. 2013. Comparative Syntax. In Carnie, Andrew, Yosuke Sato, and Daniel Siddiq. The Routledge Handbook of Syntax. Routledge.
Haugen, Jason D., & Daniel Siddiqi. 2013. Roots and the derivation. Linguistic Inquiry 44.3: 493-517.
Haugen, Jason D., and Daniel Siddiqi. 2016. Restricted Realization Theory. In Daniel Siddiqi and Heidi Harley, eds., Morphological Metatheory. Amsterdam: John Benjamins
Kaplan, Ronald. 1987. Three seductions of computational psycholinguistics. In Linguistic Theory and Computer Applications, ed. P. Whitelock et al., 149–188. London: Academic Press.
Kaplan, Ronald, and Joan Bresnan. 1982. "Lexical-Functional Grammar: A formal system for grammatical representation." The Mental Representation of Grammatical Representation.
Lieber, Rochelle. 1992. *Deconstructing morphology: Word formation in syntactic theory*. Chicago: University of Chicago Press.
Marantz, Alec. 2005. "Generative linguistics within the cognitive neuroscience of language." Linguistic Review 22, 429-455.
Marantz, Alec. 2014. Locality Domains for Contextual Allomorphy across the Interfaces. In Distributed Morphology Today: Morphemes for Morris Halle, ed. by Ora Matushansky and Alex Marantz, pp. 117-134. Cambridge, MA: MIT Press.
Newell, Heather. 2016. English Lexical Levels are not Lexical, but Phonological. Ms.
Pfau, Roland. 2009. *Grammar as a Processor.* Amsterdam: John Benjamins
Radkevich, Nina. 2010. On Location: the structure of case and adpositiona. PhD Dissertation.
Ramchand, Gillian. 2015. Allosemy—No thanks. Language Blog. 14 Sept. 2015.
Siddiqi, Daniel. 2009. *Syntax within the Word.* Amsterdam: John Benjamins.
Siddiqi, Daniel, and Heidi Harley, eds. 2016, Morphological Metatheory. Amsterdam: John Benjamins.
Swanson, Heather. 2016. Problems in replicating studies that rely on lexical frequencies. Talk given at Mo-MOT 1. Ottawa, ON.
Stump, Gregory. 2001. Inflectional Morphology: A Theory of Paradigm Structure. Cambridge: Cambridge University Press.
Stump, Gregory. 2016. Paradigms at the interface of a lexeme's syntax and semantics with its inflectional morphology. In Siddiqi and Heidi Harley (2016).