Notes to Committee on *Lexical flexibility in discourse* (Hieber)

In the following notes, one or more comments on the original draft are presented as block quotes (divided by a line separator if there are multiple comments I’ve grouped together), followed by my notes about how I chose to address each comment.

Any comments not mentioned here were minor comments which I accepted/incorporated into the dissertation as appropriate (e.g. typos, minor clarifications and changes to wording, etc.).

# Stefan Th. Gries

p. 94: 734 clauses is not 3.8% but 5.3%.

This appears to have been a miscalculation on Nakayama’s part, which I inadvertently reproduced. He does report the correct 5.3% statistic elsewhere though, so I assume this was the intended amount and updated what I report in my dissertation to say 5.3%.

p. 99-102: Calculating the dispersion values depends on how many parts the corpus has. Clarify what constitutes parts, how many parts there are, and how large the parts are for each corpus.

I added a short paragraph to the end of Section 3.4.3 clarifying that individual texts were used as corpus parts for both corpora, and the total number of texts/parts for each. The OANC has 2,410 texts/parts and the Nuuchahnulth corpus has 24 texts/parts.

p. 116: I am not sure it makes a lot of sense to report a statistical significance test for the hypotheses that the medians of the flexibility values are not 0 when Figure 4.6 and 4.7 show that 3 of the 4 medians are in fact 0.

p. 142f.: This seems to misrepresent the findings: “It was found that the overwhelming majority of English stems exhibit some degree of flexibility. That degree of flexibility is generally small (∼ 0.2), but not so small as to be statistically insignificant.” First, the statistics reported around p. 117 etc. showed clearly that the median flexibility value in 3 of the 4 samples was actually 0. Second, that’s why I above said we don’t need/want a significance test that the median is not 0.

I agree that this doesn’t make much sense to report in hindsight. I included this because it was in my original plan to look at these tests, but it felt tortured. I’ve simply removed those paragraphs, and adjusted my discussion of this portion of the data throughout the dissertation (intro, results, and conclusion) to avoid this misrepresentation.

p. 119: This discussion emphasizes how Nuuchahnulth has so much flexibility along the reference-predication axis, ok, but doesn’t the left panel of Figure 4.8 suggest a similar degree of flexibility along the reference-modification axis for English? This doesn’t seem to discussed much in comparison.

This had completely escaped my attention. I mention this fact here and elsewhere now.

Section 4.4: Resample the corpus multiple times (100? 1000?) in different (random) orders and plot the curves on top of each other to see what the "confidence interval" of the curve is.

This has been done, and the discussion around those plots updated slightly to reflect the new data.

Section 4.5: The results are probably the worst I’ve ever seen on actual data: the regression models explain extremely little. This section needs to state this more clearly. Report R2 values and explained deviances for each of the 4 models. Stating, “There is so much variability in the data for English that no firm conclusions can be drawn,” just understates the awfulness of the results. This section and p. 145 also discuss the significant findings for Nuuchahnulth in a way that inflates their predictive power.

I’ve corrected the mistakes in the code and rerun the models, all four of which now return no significant results and explain very little of the deviance. I’ve updated the discussion of this research question throughout the dissertation to clearly reflect this fact.

# Bernard Comrie

One of your conclusions is that flexibility involving modification is less frequent in Nuuchahnulth. But as you note yourself, especially towards the end of the dissertation, modification is overall less frequent in Nuuchahnulth. This leads me to wonder whether your result would have been different if you had measured not raw incidences of flexibility involving modification, but relative to the overall incidence of modification. I could imagine similar issues arising between reference and predication in languages differing in referential density.

We discussed this via email, and it was decided that I wouldn’t make this normalization, but that I would include a brief discussion of this approach and why I chose not to utilize it. That’s now in Section 3.4.1.

What criteria do you use to distinguish between compounds and phrases, in particular in English? For instance, you cite “childhood” as having flexibility between reference and modification. But what decides if, for instance, “childhood dreams” is an instance of modification within a phrase, or a compound noun? Let me hasten to add that I don’t have an answer to this, at least not one that I could defend in good conscience, but it is a potential problem, not only language-internally but also across translation equivalents.

I operationalized this in an overly simplistic but consistent way: if the term is written as a single word, I treat it as a compound; if the term is written with space between the two elements, I treat it as modification. This is obviously an imperfect operationalization. However, the orthographic convention of writing compounds as single words does at least correlate with their usage as compounds. This strategy also means that I depend on the judgments of the original transcribers of the corpus regarding compounds.

I now include the above operationalization in my discussion of the treatment of compounds in English in Section 3.2.

# William Croft

I’ve divided my responses into the same categories that you provided them—major revisions, minor revisions, and miscellaneous. For the major revisions, I’ve first provided a high-level overview of how I addressed your requests, and then made notes on some specific points from your comments. On the whole, I accepted almost all of your suggestions, major and minor, and made substantive changes to address them to the extent possible, given my timeline.

## Major revisions

### Terminology: lexical flexibility

My reasoning for retaining the term *lexical flexibility* is mainly a practical choice based on the fact that this is the widely-accepted term in the field. See the first paragraph of §2.3.1 and the last paragraph of §2.3.1.1. I agree with your criticisms of the term and echo those sentiments in various places throughout the dissertation.

I would also like my dissertation to be easily findable. A Google Scholar search for “functional expansion” returns no relevant results, whereas a search for “lexical flexibility” lists a number of highly-relevant ones. I also doubt most scholars would know what my dissertation is about if they saw the title *Functional expansion in discourse*; the current title however makes the topic immediately clear. I also want to avoid policing and/or prescribing terminology in the field *a la* Martin Haspelmath. I agree the term *lexical flexibility* is a poor one, so I make sure to precisely define how I’m using it, and then critique it while offering my alternatives (see next paragraph).

That said, while I’ve retained the term *lexical flexibility* in the title, I have made significant changes to the dissertation to pivot away from this term as soon as possible. Instead I adopt the terms *(lexical) polyfunctionality* (to describe the synchronic state of affairs wherein one form is used for multiple discourse functions) and *functional expansion* (to describe the diachronic process which gives rise to heterosemy) early in the dissertation. As a result, the term *lexical flexibility* now appears primarily in the introduction and literature review (since this is the term used by most of the previous literature).

### Methods: explanation of coding decisions + examples

Section 3.3 is now significantly expanded. In addition to including a more extensive discussion of compounding (see also my notes below), it now includes an exhaustive explanation of the coding decisions for each annotated item in the sample Nuuchahnulth passage, describing the various referring, predicating, and modifying constructions that appear. The sample annotations have been moved into the body of the text.

### Data: zero-coded vs. overt derivation

DH: ‘But to lump flexible forms in with overtly derived forms ignores the fact that there is something unique about them—namely that they can appear in different discourse functions with no overt derivational morphology.

WC: Actually, zero vs. overt coding as alternative strategies for constructions is not that unique.

But the assertion in the last sentence of the quoted passage, ‘regardless of one’s theoretical analysis of flexible forms, their behavior is substantively different from nonflexible ones’ has not been demonstrated in this dissertation. […] The only way one can demonstrate the assertion in the last sentence is to compare such words to words that occur in overtly-coded constructions expressing those same discourse functions.

I simply mean that the morphological behavior of zero-coded forms is distinct from that of overtly coded forms (in that they’re zero-coded), not that zero-coding is exceptional in some way. It was a mostly tautological statement. I am not in fact trying to demonstrate behavioral differences between overt vs. zero-coded forms in the dissertation—that would constitute an entirely separate study in itself, and is far outside the scope of this dissertation. I've clarified this in both places.

### Data: semantic shift

Indicate somehow the range of semantic shift of the words examined in the study.

I know that doing this comprehensively is impossible; it was stated somewhere in the dissertation that this data was not collected. All that can be done is some representative illustration.

For example, some zero-coded constructions across languages have a general (i.e. productive) semantic shift. […] I believe this is the case for Nuuchahnulth.

As you’ve made clear in various comments, semantic shift is a huge and complex topic. I’m wary of “representative illustration” here, because it amounts to cherry-picking examples—we can’t be certain that the examples actually *are* representative without a proper study of semantic shift, which is beyond the scope of the dissertation. I don’t want to make a claim about the productivity of semantic shifts in Nuuchahnulth with cursory data (I was hesitant to even include §4.6 on semantics). This would just add one more study to the many existing ones that present a particular pet view of lexical flexibility without backing it with clear empirical data. Moreover, based on my impression after having annotated the entire Nuuchahnulth corpus, I actually *don’t* think all or even most semantic shifts in Nuuchahnulth can be characterized in a general, productive way. But this is just my impression. I would prefer not to make claims either way in the absence of a representative/comprehensive set of annotations (not to mention a well thought-out set of coding categories for types of semantic shifts, which is arguably the more difficult task).

What I’ve done instead is expanded §4.6 to exemplify both seemingly productive and idiosyncratic cases for English and Nuuchahnulth as a way of illustrating the need for more research in this area, without making any claims about productivity either way.

### Theory: conventionalization

I’ve adjusted the framing of flexibility and conventionalization throughout the dissertation, and included a more extensive discussion of the relationship between the two, as requested in the major revisions sections of your comments.

## Specific points from major revisions

It is clear that properly addressing these issues would require major revisions of the dissertation—the empirical study itself, not just the text of the dissertation.

I’m unable to code any additional data for this dissertation in the available timeframe. The dissertation is limited to analyzing the data already collected and the annotations already made on those data.

This description as *lexical* implies that this is fundamentally a phenomenon about *words* or *word classes*. […] I think that the dissertation hasn’t really abandoned a word-class-based approach to the degree that it should, which leads to some confusing and/or contradictory statements.

Words, yes. Word classes, no. This dissertation doesn’t attempt to make generalizations about word classes. Its aim is more modest. It merely describes the range of zero-coded discourse functions for individual lexical items, and in the aggregate. I’ve adjusted the prose in places to make this clearer.

The examples you provide where the dissertation doesn’t sufficiently “abandon a word-class-based approach” are cases where I am simply reporting previous research, sometimes retaining their terminology. As you noted, it is often difficult to translate previous research into a constructional framework. Nonetheless, I’ve updated the prose in these and other cases in an attempt to be more precise.

On p. 120, Nakayama is quoted to say that he ‘treats “adjectivals” as a subclass of verbs’ in Nuuchahnulth, without comment. Yet this is problematic in at least two respects. First, the use of terms like ‘adjectival’ and ‘verb’ assumes that language-specific word classes are comparative concepts—a position that was argued against in chapter 2. Second, ‘class’ vs. ‘subclass’ employs methodological opportunism, something which was also argued against (specifically with respect to class vs. subclass) in chapter 2.

This is a case where I was reporting results in the author’s own terminology. I’ve revised the summary of this research to avoid these terminological problems.

DH: ‘In a corpus analysis of English and Mandarin, Thompson finds that property words have primarily two functions in discourse: to introduce new discourse-manipulable referents, and to predicate attributes of an already-known referent. In English these two functions are realized via attributive adjectives and predicative adjectives respectively.’

WC: Not really: these two functions are expressed by attributive (modifier) constructions and predication constructions respectively. These are not properties of words; these are properties of constructions.

This is another case where I report results in the author’s own terminology. I’ve revised the wording here.

Lexical flexibility is defined as ‘the use of a lexical item in more than one discourse function (reference, predication, or modification) with no overt derivational morphology’ (139). The definition is not quite accurate as to what the dissertation actually does. […] Referring only to derivational morphology in the definition is too restrictive[.] […] The more general notion is *overt coding* […] as opposed to *zero coding*—two different strategies for expressing the function. A more accurate definition of “lexical flexibility” is ‘the use of a lexical item in *constructions* expressing more than one discourse function and employing a *zero coding strategy*’.

This definition is more in line with what I intended, so I’ve adopted a similar version and updated the prose where relevant.

To me, the statement that ‘Nuuchahnulth does not have dedicated modifying constructions’ is incorrect (see also p134). Nuuchahnulth’s modifying construction was described just above: ‘a conventionalized syntactic construction in which the modifier precedes its head’. That is a dedicated morphosyntactic construction. The fact that there is no morphological inflection or derivation doesn’t mean it’s not a construction—at most, it is confusing the discourse function of modification with associated constructions such as degree, indexation, etc. This statement also appears to be confusing the existence of a construction—a morphosyntactic structure that expresses a function—and a strategy used to encode that function. That is, it appears to assume that the use of a zero strategy of a function represents the absence of a construction for that function.

I’ve adjusted the wording and use of these terms here and elsewhere. I’ve also cited your relevant works in a clarifying footnote the first time the word *strategy* appears in the main body of the text.

The quoted passage also leaves aside the strategy of “compounding”, in the morphophonological sense of a phonologically bound combination of two morphemes, which plays a significant role in Nuuchahnulth.

Here is an example of what sort of issues need to be explained to the reader, one that can be inferred from the small number of examples in the current draft. On pp. 92-93, it is said that so-called “compounds” are excluded, that is, *back* in *back yard, hard back book* and *back burner* were excluded. But in Appendix C, *street* in *street clothes* is included, and coded as a modifier (165). How are “compounds” and “nouns functioning as modifiers” differentiated? Can they really be differentiated?

I state in §3.3.1 that “Compound words were included in the analysis, but individual components of compound words were not.” Compounds are treated as a lexical unit for the purpose of this analysis.

Also, despite their superficial similarity to each other, Nuuchahnulth Lexical Affixes are not a type of compounding. Nuuchahnulth does not have a compounding strategy (with the possible exception of a “Nominal Concatenation” strategy [Nakayama 2001: 90], which is quite rare). Nakayama (2001: 18) provides good evidence that Nuuchahnulth Lexical Affixes are indeed affixes rather than compounding / noun incorporation. I now summarize Nakayama’s evidence for the fact that Nuuchahnulth Lexical Affixes are not elements of compounds in §3.3.1.

Section 3.3.1 now also includes a more extensive discussion of compounds, how they are identified in each language, and how they are treated for the purpose of analysis.

I hate to say this, but one could argue that the dissertation could also be accused of methodological opportunism, since it dismisses or at least sets aside semantic shift, which seems to be treated as not important to the goals of the dissertation.

Obviously, I think this statement is incorrect. If the dissertation were trying to establish a theory of parts of speech, or argue for the existence of large word classes, or argue that flexible lexemes are truly “flexible” / productive, then I do think this criticism would be appropriate. However, this dissertation has more modest aims. It investigates the set of discourse functions that a given linguistic form can be used for (as noted by your next comment below). It is an investigation into the *extent* of flexibility, not the semantic behavior of flexible forms. It makes no claims as to their degree of relatedness or their lexical unity. I agree wholeheartedly that the semantic shift involved in cases of functional expansion is deeply interesting, but it is not the *only* interesting facet of these cases, and it is not the facet that I examine in the dissertation.

I was surprised to see that many of the “properties” are numerals and quantifiers, in both the examples and Table 4.4 (see also p 21). Table 4.4 lists a wide variety of semantic types of modifiers, only a minority of which are usually called properties […] I would definitely not include numerals or quantifiers as “prototypical modifiers”, or even ‘proper’ (example 38a, page 135).

I’ve updated the description here to avoid calling these properties / prototypical modifiers.

## Minor revisions

p2, fn 1: this is too important to put in a footnote; integrate into the main text.

I’ve opted to leave this as a footnote. Placing it in the main body of the text is too disruptive to the flow of the prose here in the very first paragraph of the dissertation, where I need to make a concise statement of the problem and engage the reader’s interest.

p7, displaying lexical flexibility: didn’t Evans and Osada (2005) also say this?

Not explicitly that I could tell (assuming you’re referring to my statement, “Lexical flexibility is not as rare or marginal as traditional approaches to word classes lead one to believe.”). They just note a few high-profile cases.

p20, bottom: actually, modifiers in general are infrequent in discourse compared to referent expressions and predicates; it’s not just Nuuchahnulth.

Also, this is probably the point where I should ask, how does the fact that object predication and action reference are so much rarer than action predication and object reference (see Croft 1991) affect the quantitative results?

Bernard asked this same question, prompting me to add a paragraph about it in Section 3.4.1. I’ve also added a note about this difference in frequency in the discussion on p. 20.

p37: if you prefer *functional expansion*, why don’t you use it? I give a lot of reasons above why *lexical flexibility* is a bad description of the phenomenon.

See my notes on major revision #1. Also note that I provide an explanation of my choice in that same paragraph.

§2.3.2 [issues regarding the “Key Findings” section]

But some of the “key findings” don’t make any sense in a constructional approach, so we don’t know what the real empirical generalizations are, or even whether they are real empirical generalizations.

This section is not necessarily an endorsement of the findings of previous authors, just a summary of themes in previous relevant research. I criticize previous research in §2.3.3: Problems & Critiques.

Also note the following in the chapter’s introduction:

“Section 2.3 summarizes the key concepts and findings that have arisen from the research on lexical flexibility. Such research, however, is not without its own shortcomings. Section 2.3 also presents the main criticisms that have been leveled against flexible analyses of word classes.”

To make this even clearer, I’ve reframed that section as “Themes in previous research” instead. I also now clearly state that I am not necessarily endorsing this research, just reporting it and then critiquing it in the following section.

Regarding “level” for flexibility (§2.3.2.3), the “key finding” is just that it varies.

Yes. One might consider this a trivial finding, but it’s one that persistently arises in the literature, and the point of this section is to give a synopsis of existing research.

item-specificity (§2.3.2.4) is another way to say that there are no generalizations, that is, there is arbitrariness

Item-specificity does not necessarily entail an inability to make generalizations. The principle of semantic shift is a counterexample to this. The meanings of individual lexical items in different discourse contexts are language-specific and item-specific and thus arbitrary/unpredictable, but they nonetheless still adhere to crosslinguistic universals regarding semantic shift.

[item-specificity is] a good antidote to some of the more simplistic theories about “lexical flexibility”

That’s why I discuss item-specificity here—it lays the foundation for a later criticism of literal/naive interpretations of “lexical flexibility”.

p75, ‘semantically more complex’ - I don’t consider this a characteristic of typological markedness, and I argue against it in *Typology and Universals*

See our email exchange about this. “semantically more complex” was the wrong phrasing to use here; my intention was to state that semantic *shifts* are universal.

p81, ‘The definition of lexical flexibility given here allows for any degree of semantic shift. Croft admits this possibility explicitly: “It of course a priori possible to construct a typological classification of parts-of-speech systems using only structural coding and allowing any degree of semantic shift.” (Croft 2001b: 68)’

but I don’t think it’s *right*

But note my next sentence: “Of course, I am not concerned here with constructing a classification of parts of speech—quite the opposite, in fact.” I then proceed to explain that the definition of lexical flexibility I’m presenting is designed specifically *not* to rely on any notion of parts of speech. This dissertation simply examines the range of functions that individual lexical items appear in, without making sweeping generalizations about how those items are grouped into classes or their functional behaviors (which would indeed require attentiveness to semantic shifts). This dissertation is much more modest in its aims. I’ve adjusted the wording here to try to make this clearer.

p116, ‘When lexical items in English and Nuuchahnulth exhibit flexibility, it is typically not to a marginal degree’ - this strongly implies semantically distant, distinct uses or senses—conventionalization, rather than “flexibility”

I’ve adjusted my analysis / presentation of the English vs. Nuuchahnulth data, so this note isn’t quite relevant anymore. Also, a high flexibility rating *per se* isn’t necessarily suggestive of distinct senses. I also discuss conventionalization in more depth now (see my notes on major revision #5 above).

p118, Figure 4.9, Nuuchahnulth - if more than half of Nuuchahnulth words are inflexible, why aren’t there most of the dots piled up on the corners of the triangle? (or at least the reference and predication corners)

Items with a flexibility of zero occur exactly in the corners, and thus overlap. These ternary plots are really more for visualizing the functions of items which *do* exhibit some flexibility. I’ve clarified this in the text.

On page 119, you say Nuuchahnulth displays ‘rampant flexibility along the reference-predication axis’. How rampant is it if more than half of Nuuchahnulth words are inflexible?

I toned down this wording.

p122 - it might be useful to compare Nuuchahnulth with an English plot of similar corpus size, to see how similar/different the plots are for the two languages

I needed to significantly adjust the visual presentation of these data here for other reasons. The new plots make this comparison easier. Note that the point of these plots is to show how flexibility changes as the size of the corpus grows, so you can easily see how English behaves at the smaller corpus sizes like Nuuchahnulth.

p128, interaction of high frequency and flexibility: is this because higher-frequency items are more likely to occur in non-prototypical functions at all in such a small sample (see second point under p118)?

After fixing some issues with the model, there turned out to be no significant interactions, so I decided not to speculate on causes / correlations here.

p129, top - good, but I fear that this shows that some of the results for Nuuchahnulth are due to small sample size rather than some linguistically interesting difference

This is only partially true. As I explain and provide plots to illustrate in §4.4, the majority of words in the Nuuchahnulth corpus don’t occur enough times to get a clear assessment of whether they've been conventionalized in multiple uses. However, for those that do occur frequently enough, there is a strong tendency for the word to have multiple discourse functions. I’ve tried to make this point more prominent.

p131, Table 4.3 - this would be more useful if examples were given of sentences with each use

Unfortunately this isn’t practical for this table due to formatting considerations. However, I now include examples for each of the attested functions of each of the items in the 100-item samples (which include the items in Table 4.3) in Appendix B.

p132, end of §4.6.1 - it could be that the “highest-flexibility words” are word forms with discrete senses in each discourse function, and so word-form frequency is really the sum of multiple word sense frequencies.

Yes, this is a definite possibility. It is also possible that the expansion into new discourse contexts is offset by a decrease in frequency for the original function, or that this varies by lexical item. Unfortunately, the data in the dissertation are inconclusive in this regard. I’ve updated the prose in §2.5 and §4.5 to discuss this possibility in more depth. It was also already discussed in §1.3 and §5.2.

p134, ex 37a-b - did you exclude these cases from your count? One could argue that these are instances of zero-coded, albeit phonologically bound, modification. But this should be discussed in your coding criteria, as I recommended.

I discuss this in the section on coding criteria now (§3.3.1), with a number of examples. In short: no; property roots with a lexical affix are excluded from the analysis. My analysis looked only at stems. Cases like ‘two-canoe’ were treated as a single lexical unit for the purpose of the analysis.

p144, 2nd paragraph - as I wrote above, it is kind of contradictory to talk about “conventionalized flexibility”.

See my notes on major revision #5.

p146, ‘When property words are used to introduce a new referent into the discourse, referring constructions are used (usually with the definite suffix *-ʔiː*)’ – this is no different from English; modifier constructions are modifiers in referring expressions

This discussion has been revised in line with your comment.

p148, middle: Lexemes should also be described in terms of the semantic shifts they undergo in the different contexts, including the possibility of divergence/split

I added a clause about semantic shifts to the relevant sentence.

p148, ‘Speakers know the range of contexts that a given form may appear in and use it that way’ - this is probably due largely to their knowledge of the semantic shifts of the lexeme

I added a note about semantic shift here as well.

pp150-51, looking at other languages: actually, it is *less* interesting to look at a theoretically biased selection of languages such as this list, rather than a stratified sample of languages that are likely to represent a broader range of types

This list of languages is aimed at empirically confirming/disconfirming previous claims about these languages. Much has been written about their flexibility, but none of these claims should be considered empirically validated. Applying the techniques in this dissertation to those languages would have the benefit of empirically confirming or disconfirming those claims. I’ve updated the prose here to also mention using a balanced sample though.

p152, looking at “flexibility” of corpora over time: actually, it would be better to look at the evolution of discourse function constructions over time, or derivational constructions over time. This would of course also involve looking at the words occurring in those constructions over time

Both approaches would be valuable.

p153 - it’s not about a language developing dedicated constructions for different discourse functions; it’s about developing multiple constructions for the same discourse function, some using zero coding and others using overt coding.

Both are relevant. I’ve expanded the text here to reflect this.

Appendix A - this is short enough to go into the text (though I’d keep the timeline)

I moved this appendix to the main body of the text.

Appendix C - as noted above, these examples all belong in the text

See my notes on major revision #2.

## Typos etc.

p70, ‘semantic space’ - I called this a ‘conceptual space’ in *RCG*, and the term ‘semantic space’ is a lot older than Finch 2003 (admittedly, a dictionary)

I added a reference to *RCG* and the term ‘conceptual space’.

p51, bottom: I don’t understand why ‘no such metaphor exists for verbalizations’

I clarified the prose here.

All other notes in the *Typos etc.* section were addressed.