The nature of the phonology-syntax interface, from variable adjective and noun word order in Tagalog

Stephanie S Shih  
University of Southern California  
shih@usc.edu

Kie Zuraw  
University of California, Los Angeles  
kie@ucla.edu

November 2018

This manuscript is material from an earlier, unpublished version of Shih & Zuraw 2017 (in Language: Phonological Analysis) that did not appear in the published version after the final round of revisions, which requested that the paper shift its focus away from theoretical consequences for phonology-syntax interface design. Because the following discussion has been requested subsequently from readers of Shih & Zuraw 2017, we are making the material available. In citation, please reference both Shih & Zuraw 2017 and this manuscript, as follows (please add lingbuzz link):

Shih, Stephanie S & Kie Zuraw. 2018. The nature of the phonology-syntax interface, from variable adjective and noun word order in Tagalog. Unpublished manuscript, USC & UCLA.

For full details of the phenomenon of adjective-noun word order variation in Tagalog, please see Shih & Zuraw 2017.

1 PHONOLOGY-MORPHOSYNTAX INTERACTION

In Shih & Zuraw 2017, we present a corpus study finding that phonology affects word order. Following previous work on phonology-syntax interaction (e.g., Zwicky & Pullum 1986a:80), we consider word order and linearization principles to be the purview of (morpho)syntax. Meanwhile, we consider allomorphy to be the purview of (morpho)phonology. As shown above, the Tagalog linker morpheme has two allomorphs (–ng and na), depending on what sound the preceding word ends with; allomorph choice therefore cannot take place until segmental phonological information is available. The study here thus engages several major theoretical issues at stake for the phonology-morphosyntax interface, which we outline in this section. To preface our conclusions, we find that phonological information—including surface phonological outputs of allomorphy—conditions adjective/noun order variation.

---

1 An alternative view is that syntactic structure gets translated into prosodic structure, and prosodic constituents are linearised by phonology (i.e., PF Movement; see e.g., Agbayani & Golston 2010). However, given evidence presented in this paper, some morphosyntactic information—i.e., lexical category specifications of noun and adjective—must still be maintained, since there is a default preference for the order adjective+noun.

2 These questions are intricately interconnected, but for clarity, we have attempted here to separate out the individual issues and their predictions for Tagalog as much as possible.
The first issue we take up is the locus of word order choice and variation in the linguistic system. That is, is word order variation located in the grammar or is it extra-grammatical? Our data show that phonological constraints that are active elsewhere in the morphophonological grammar of Tagalog also play a role in adjective/noun order variation, suggesting that such variation should be located in the grammar itself (see e.g. MacKenzie 2012:273ff for a similar argument on how to determine the locus of variation). Second, what is the nature of the phonology-morpho-syntax interface? Our results point to a view of the interface in which phonology can interfere in word order, and in which surface phonological information must be available at the point of linearization choices by the morphosyntax. Finally, we examine here which phonological conditions are likely to affect word order and suggest that the phonological conditions we should expect are those that are active elsewhere in the grammar.

The nature of the syntax-morphology-phonology interface is a central issue for theories of the linguistic system, and cross-cuts theory-internal architectural debates. Given the wide-ranging generality of the subject, we choose to address the interface in this paper at a high level, remaining agnostic about the specifics of particular theoretic implementations. In doing so, we follow a precedent in the literature. For example, Zwicky & Pullum remark (1986a:73): “the positions we are taking have a very high degree of generality and applicability; if they can be convincingly defended, then they will have consequences for any imaginable theory of syntax” (see also Miller et al. 1997:68 for discussion of this point).

Ultimately, the view of the linguistic system that our results point to is one in which phonology and morphosyntax interact in the grammar—in particular, for linearization and word order. This view of an interactive phonology-morphosyntax model is schematized in a general fashion in (1). In (1), the first linearization produced by syntax is adjective+noun order, which is the preferred order overall, as we will show from our data (§ Error! Reference source not found.). The phonology cashes out this word order into a phonological form, including choice of linker allomorph, and checks the result for well-formedness: in the example here, itim na peluka violates an Obligatory Contour Principle (OCP) restriction against adjacent nasals (mn) (see § Error! Reference source not found.). Then, the information returns to the syntax for an alternative linearization if the first option is deemed insufficiently phonologically well-formed, as in the illustrated case. The process from syntax to phonology is repeated, and a word order that satisfies phonological conditions—noun+adjective, pelukang itim, which does not have a nasal OCP violation—is the ultimate output of the grammar.

---

3 Of course, there may be certain types of variation that exist outside the grammar (e.g., sociolinguistic conditions). The debate of grammar versus use extends beyond the narrower focus of this paper, which is on phonological conditions of word order variation.

4 We remain agnostic here about whether this interaction exists between entire modules of phonology and syntax proper in the grammar, or whether the interaction exists within a subsection of the grammar—e.g., as in spell-out/Phonological Form, where linearization occurs (see e.g., Embick 2007; Tucker & Henderson 2010; Tucker 2011 for similar Distributed Morphology-based proposals).
The schematic in (1) shows a derivational version of a phonology-syntax interaction, but this can be easily translated into a parallel Optimality-theoretic interaction in which some phonological constraints (P) outrank (⇒) or outweigh morphosyntactic constraints (S): i.e., S ⇒ P is not always true (see also Harford & Demuth 1999; Teeple 2008; cf. Golston 1995). The schematic is also relatable to a psycholinguistic model in which syntactic and phonological activations can overlap, instead of existing strictly successively (cf. e.g., Bock & Levelt 1994).

Our conclusions run contra the dominant assumption in linguistic theory that information flows unidirectionally between parts of language: that is, that (morpho)syntax occurs prior to and with no interference from phonology. Zwicky & Pullum (1986a, 1986b; Miller et al. 1997) term this one-way separation the ‘Principle of Phonology-free Syntax’; see also ‘The Separatist Hypothesis’ and other manifestations of the same idea throughout linguistic theory (e.g., Beard 1966). Under the strongly separatist, strictly unidirectional view of the interface, phonological information should have no import to word order, and word order is predicted to be determined solely by morphosyntactic principles. We will show that this is incorrect: phonological factors do condition word order in our data (see also e.g., Wolf 2008; Shih 2014, 2017 for surveys of such effects).

In some versions of the unidirectional view, word order variation is relegated to extra-grammatical processes, with an extra-grammatical ‘filter’ in place to produce output variation (e.g., Zwicky & Pullum 1986b; Miller et al. 1997). The phonological factors conditioning word-order variation in our data are active elsewhere in the grammar, and it is unmotivated to duplicate grammatical principles in an extragrammatical component.

Other versions of the unidirectional view allow for limited phonological conditioning by employing phonology as a filter on equally grammatical outputs of the morphosyntax. This is often instantiated in Optimality-theoretic terms as an a priori ranking of syntactic constraints (S) over phonological constraints (P), S ⇒ P (e.g., Golston 1995; Anttila 2016). Treating phonology as a filter makes two predictions that are not borne out in our data. First, absent any relevant phonological factors, the two word order options should be equally grammatical in this filter model.⁵ We

---

⁵ assuming that the syntactic component is non-gradient.
find, however, that one order is strongly preferred. More crucially, the phonological filter should apply equally to all word order options provided by the syntax in this model, but we find that phonological conditions apply more strongly to one order.

One major factor that has led to the dominance of phonology-free syntax in language models to date is the apparent rarity of phonological conditioning on morphosyntactic behavior. The classically-cited example of a phonology-syntax interaction that should not exist, from Zwicky & Pullum (1986a:75, their emphasis), is “a movement transformation that obligatorily moves ... [a] constituent that begins phonetically with a bilabial consonant,” maintaining that surface segmental information should not affect word order (see also arguments that prosody but not segmental information can affect syntactic behaviors: e.g., Nespor & Vogel 1986; Inkelas 1990; Zec & Inkelas 1990). We argue that the reason we don’t observe bilabial fronting is not because of grammatical architecture, but because bilabial-induced movement is unlikely to be preferred by a phonology. In the view of grammar we argue for here, a similar movement rule could occur if it were motivated. For example, suppose that there was a condition, active in the phonology of the language, against two labials in succession (an instance of the Obligatory Contour Principle, Leben 1973). Then we would expect syntactic repairs such as fronting of e.g., [am] [ba] to [ba] [am] (see also Shih 2014:213ff).

The table in (2) provides a summary of the questions and predictions that different conceptions of the syntax-phonology interface have for Tagalog adjective/noun word order. The conclusions argued for in this paper are highlighted in bold. Further detailed discussion and comparison of phonology-syntax interface models is taken up in §2 after the results of the study have been introduced.

<table>
<thead>
<tr>
<th>Interface issue</th>
<th>Predictions for Tagalog</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Locus of word order choice?</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Word order variation occurs in the grammar.</strong></td>
<td>Factors that condition word order will be ones found elsewhere in the Tagalog grammar.</td>
</tr>
<tr>
<td>vs. Word order variation is extragrammatical.</td>
<td>Factors that condition word order should not be grammatical ones.</td>
</tr>
<tr>
<td><strong>Nature of the interface?</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Phonology affects word order.</strong></td>
<td>Phonological and morphosyntactic factors condition word order. Surface phonological patterns are optimized via word order alternations; phonology can temper a morphosyntactic preference; and phonological factors can apply differentially to word orders.</td>
</tr>
<tr>
<td>vs. Word order is blind to phonology.</td>
<td>Only morphosyntactic factors condition word order.</td>
</tr>
</tbody>
</table>

---

6 Moving a constituent with a bilabial consonant, as in Zwicky & Pullum’s example, wouldn’t repair a simple restriction against labial consonants, but if the restriction is against labial consonants in some sort of syntagmatic configuration, such as *[labial][labial]*, then changing the word order can fix the problem, by removing the illicit sequence of [labial] consonants. In this way, word order alternatives from the syntax can be candidates to phonological optimization just as phonological repairs are.
vs. Word order options are filtered by phonology. Phonology can only choose among equally grammatical word order options, and must apply equally to all options.

Which phonological effects?

Segmental phonology affects word order

Segmental phonological information such as contextual markedness or repetition avoidance condition word order.

Phonological conditioning is not random

Phonological factors that condition word order are found elsewhere in the grammar.

2 DISCUSSION

The investigation presented in Shih & Zuraw 2017 examined how phonology influences word order variation in Tagalog. Our corpus study revealed that, controlling for non-phonological predictors, surface phonological information and phonological well-formedness conditions significantly contribute to predicting variable adjective and noun order, an alternation that had previously been characterized as primarily “free.”

With results in hand, we now return to the theoretical issues for the phonology-syntax interface laid out in §1: where is the locus of word order choice in the linguistic system (§2.1)?, what is the nature of the phonology-morphosyntax interface (§2.2)?, and which phonological conditions interact with morphosyntactic behaviors (§2.3)? Our results support a model of the interface in which phonology and syntax interact in the grammar, both phonological and morphosyntactic conditions affect word order, and surface segmental information from phonologically-conditioned allomorphy and the phonological grammar must be available for adjective/noun ordering in Tagalog, as shown in (3), reproduced from (1):

(3)=(1) Interactive phonology-syntax model

![Diagram of Interactive phonology-syntax model]

Syntactic analysis and the phonological grammar: pelukang itim

Syntax

if *Adj+Noun, try Noun+Adj

Phonology

*itim na peluka

Adj+Noun

OUTPUT

pelukang itim
2.1 Grammatical locus of word order variation

The first issue that we take up here is where Tagalog adjective/noun variation is located in the grammar. This is a crucial question because proponents of Phonology-free Syntax relegate variable phenomena to outside the "core grammar"; thus, under such a treatment, Tagalog adjective/noun variation would be considered language use (not grammar) and be conditioned by an extra-grammatical filter component that chooses between word order outputs provided by the grammar (Zwicky & Pullum 1986a, 1986b; Miller et al. 1997). As schematized in (4), under this view, syntax entirely precedes phonology, producing both adjective+noun and noun+adjective orders. Phonology then spells out each of the syntactic orders, and the grammar produces two grammatical outputs. In the language use component, an extra-grammatical filter chooses between the adjective+noun or noun+adjective options: here, itim na peluka is penalized because it has two adjacent nasal segments, which the extra-grammatical filter disfavors.

(4) Extra-grammatical filter model

Indeed it is the case that much of the existing evidence of phonological interference in morphosyntactic processes comes from variable sources, and this has fed the argument that purported
cases of phonology affecting syntax are extra-grammatical, a fact of language *use*, and thus irrelevant to the architecture of grammar.

However, many have argued that categoricity should not be the diagnostic of whether something belongs in the grammar. Further, the grammar itself—depending on its design—can produce non-categorical outputs, using the same grammatical principles that derive categorical patterns: e.g., ‘comparative grammaticality’. Such an approach that derives gradiency in the grammar has its roots in rule-based theories (e.g., Labov 1969) and has been increasingly embraced across linguistic subfields: in phonology (e.g., Anttila 1997, 2002; Boersma 1998; Zuraw 2000; Goldwater & Johnson 2003; Coetzee & Pater 2011; Coetzee 2014), morphology (e.g., Hay & Baayen 2005), syntax (e.g., Legendre 2001; Manning 2003; Bresnan et al. 2007; Sag & Wasow 2011), sociolinguistics (e.g., Labov 1972; MacKenzie 2012), and usage-based approaches to language (e.g., Bybee 2001). Under Occam’s Razor, there should be no need to duplicate principles, if we find that they seemingly occur both in variable phenomena and in clear grammatical cases (e.g., Guy 1997; MacKenzie 2012:273ff).

Our conclusions from Tagalog adjective/noun ordering clearly illustrate that the same phonological principles that drive morphophonological alternations and phonotactic patterns in Tagalog also affect word order variation; therefore, the locus of phonological interference with word order variation should be in the grammar. Take, for example, the avoidance of nasal-voiceless consonant sequences (*N̥C*) in adjective/noun ordering. As discussed in Shih & Zuraw 2017, adjective/noun orders that avoid illicit N̥C sequences are more likely to occur. We find the *N̥C* condition at work within the morphophonological grammar of Tagalog as well, at prefix-stem boundaries (Zuraw 2010). In order to avoid N̥C sequences, stem-initial voiceless obstruents will fuse with preceding nasals: e.g., /ma-paŋ-kamkám/ → [ma-paŋamkám], ‘rapacious’. Zuraw (2010) shows that, despite the variation and lexical specificity that the constraint demonstrates in Tagalog, *N̥C* is productive for nonce words and loanwords, and should therefore be located squarely within speakers’ grammars (versus experiential knowledge of language use) (see also Pater 1996, 2001 for more on *N̥C* cross-linguistically).

Other phonological conditions in our study that also have effects elsewhere in the Tagalog grammar include nasal OCP (i.e., repetition avoidance) and morphophonological alignment and syllable structure optimization. Since the same conditions reported elsewhere in the grammar are found to predict adjective/noun ordering, we maintain that the locus of phonology-syntax effects in ordering variation is not in a post-grammatical filter that duplicates grammatical functions but rather squarely within the grammar itself.8

2.2 Nature of Phonology-Syntax Interface

Having established that phonological effects on word order variation in Tagalog should be located in the grammar, we now turn to the question of how phonological conditions interface with the

---

7 MacKenzie still allows for a probabilistic module of language use that is separate from the grammar, for constraints that make up ‘speech style and usage frequency’ as well as psycholinguistic conditions (2012:280; see also Coetzee & Kawahara 2013). We remain agnostic here on the possibility of a separate module for language use and what the architecture of such a module is. Our goal is simply to locate the place of phonological effects on word order variation, for which our results suggest a grammar-internal analysis.

8 That said, it is worth noting that even if all variation were considered the domain of language use, one still needs a model of language use to explain the variation that is observed. Here, too, the same questions of how phonology and syntax interact arise: see e.g. Vigliocco & Hartsuiker 2002 for a summary.
process of deciding word order: specifically, at what point in the grammatical system do the phonological conditioning factors identified here influence adjective/noun ordering?\footnote{The majority of literature on the phonology-syntax interface has focused largely on the opposite question: namely, how much, and what kind of, syntactic information gets passed to the phonology? While the possibility for the flow of information from syntax to phonology is fairly well-established by the bulk of this existing literature, the possibility of the flow of information from phonology to syntax—though not as widely discussed—is much more controversial in comparison.}

We consider here two conceptions of the morphosyntax-phonology interface, to represent the ends of the spectrum of extant views on the matter: in one view, syntax is blind to phonology, and phonology merely acts as a filter to syntactic outputs; in the other view, which we endorse (see (3)=(1)), phonology interacts with syntax in conditioning linearization and word order. As in §1, we maintain a high-level discussion here, and the conceptions of the morphosyntax-phonology interface discussed are meant to encompass basic ideas that may be cashed out in specific theories with some variation.

One view, which we call the phonology-as-filter model, maintains that phonology has no interference in morphosyntactic processes. Under this view, as schematized in (5), both adjective/noun orders are generated as viable outcomes by the (morpho)syntax and are then compared for phonological well-formedness. The output of the grammar is the phonologically-filtered word order, \textit{pelukang itim} (\textit{itim na peluka} is ruled out by the phonology on the basis of a nasal OCP violation).

\begin{itemize}
\item (Morpho)syntax
\item \hbox{Adj+Noun, Noun+Adj}
\item Phonology
\item \hbox{*itim na peluka, \checkmark pelukang itim}
\item OUTPUT
\item pelukang itim
\end{itemize}

\begin{center}
\begin{tabular}{c}
\hline
\textbf{OUTPUT} & \textbf{pelukang itim} \\
\end{tabular}
\end{center}

The phonology-as-filter model, which eschews phonological information in syntax, is akin to Phonology-free Syntax, but with variation modeled by a portion of the grammar. In parallel Optimality-theoretic (OT) terms, the phonology-as-filter model is captured by the \textit{a priori} and categorical ranking of syntactic constraints (S) above phonological constraints (P), \( S \gg P \) (e.g., Golston 1995\footnote{Golston 1995 argues, however, that morphological constraints are all ranked \textit{below} phonological ones.}; Anttila 2016\footnote{Anttila 2016, however, provides no ranking argument demonstrating that syntax \textit{must} be ranked above phonology in the tableaux provided.}). Golston (1995) makes the argument that because both syntactic options are equally
grammatical, phonology must do the filtering after syntax has produced grammatical options. Relatedly, arguments have been made in Distributed Morphology (DM), where the assumptions of Feature Disjointedness and Late Insertion are that phonological properties are not present until post-syntax (e.g., Marantz 1995; Embick 2000:187–188; see also ‘The Separatist Hypothesis’, Beard 1966): this view seems to imply that the phonological component has the job of choosing between word orders. In psycholinguistic terms, this would meant that grammatical encoding (i.e., encoding of syntactic structures) and phonological encoding are independent processes that experience only unidirectional and sequential information flow from the grammatical to phonological modules (e.g., Levelt 1989; Bock & Levelt 1994; Ferreira & Slevc 2007).

The specific predictions that a phonology-as-filter model makes for Tagalog are that the two alternative orders should be more or less equally grammatical (assuming a categorical syntactic component), and crucially, that the two word orders should compete within the phonology equally, without reference to their morphosyntactic identities (that is, phonological conditions should apply equally to the adjective+noun and noun+adjective orders). Some forms of the phonology-as-filter model (e.g., the Late Insertion approach) also predict that the surface phonological form of the linker allomorph should not affect linearization and word order, since linearization should be completed before the insertion of segmental phonological information (see e.g., Embick & Noyer 2001, where linearization occurs several steps before phonological form in the PHONOLOGICAL FORM branch).

The other view of the interface, which we endorse here, is one in which phonological considerations interact with morphosyntactic behaviour, particularly in linearization. This is the interactive phonology-syntax model. A derivational characterization of the interaction, as shown in (3)=(1), is that one, more-basic word order (adjective+noun) is produced by the syntax as the default and considered by the phonology for well-formedness, which then requests a morphosyntactic alternative from the syntax if the default is found to be phonologically ill-formed. This sort of iterative loop between morphosyntax and phonology through the linearization process aligns with some accounts of linearization proposed in, for example, DM, where the interleaving of linearization and phonological processes has at least been posited for prosodic structure-building (e.g., Tucker & Henderson 2010; Tucker 2011), or as multiple cycles through the PHONOLOGICAL FORM branch of the grammatical model (e.g., Embick 2007). Under an Optimality-theoretic approach, the interactive phonology-syntax model is captured by the potential for phonological constraints to outrank syntactic ones, P » S: Harford & Demuth (1999), for example, argues for P » S by demonstrating two opposing patterns from related languages, and Teeple (2008) argues for the existence of P » S within a single language, contra Golston’s (1995) claims. In psycholinguistic terms, the interleaved approach calls for bidirectional information flow—or at least, overlapping information flow—between grammatical and phonological encoding (e.g., Vigliocco & Hartsuiker 2002).

Evidence from Tagalog adjective/noun ordering suggests that phonology participates in word order, providing evidence for the interactive phonology-syntax model and against the phonology-as-filter model. First, the surface form of the linker morpheme plays a role in predicting adjective/noun order. The –ng form of the linker is preferred to the na form. Additionally, the surface form of the linker is used in assessing the phonological conditions of nasal OCP, contextual markedness, and hiatus avoidance—all of which are shown to significantly predict adjective/noun order. This result demonstrates that the surface phonological form that is the output of morphophonological alternations must be available during word order choice (i.e., linearization).
Second, our corpus investigation suggests that the two possible orders of adjectives and nouns in Tagalog are not equal options, as a phonology-as-filter model (with a categorical syntax that outputs equal grammatical options) assumes. Adjective+noun order is overwhelmingly preferred, even once controls for individual words and frequencies are included. This quantitative result aligns with observations in the previous Tagalog literature that one order is more basic—or, default—than the other.

Third and most crucially, we find that the conditioning factors with the strongest effect sizes tend to be the ones that regulate phonological well-formedness in the adjective+noun order—that is, phonological ill-formedness in the default order appears to be more strictly penalized, triggering a deviation from that order. For example, although the surface form of the linker for both adjectives and nouns contributes to order choice, the linker form of the adjective has a stronger effect (adj. linker $\beta=1.2793$, vs. noun linker $\beta=-0.7744$). Similarly, nasal-nasal OCP avoidance is enforced more strongly when a prenominal adjective (i.e., adj in adjective+noun) is nasal-final ($\beta=1.224$) than when a nasal-initial adjective is post-nominal (i.e., adj in noun+adjective), following an –ng linker ($\beta=0.224$). The results also show a stronger effect of *NC in the default adjective+noun order, where a voiceless consonant-initial noun would form an illicit NC cluster with an –ng linker on the adjective (adjective+noun $\beta=-0.3299$, vs. noun+adjective $\beta=0.1992$). One possible exception to this otherwise strong pattern of adjective+noun order primacy is alignment and hiatus avoidance, for which our model shows an effect for vowel-initial adjectives in noun+adjective order but not vowel-initial nouns in adjective+noun order.

The model of adjective/noun word order choice reported here points to a preference for one order, which is more frequent overall. There are more and stronger active phonological (and non-phonological) constraints that penalize phonological ill-formedness in that default order. We take these results as evidence for an interactive phonology-syntax model, in which some amount of phonological surface information (that has undergone morphophonological operations) feeds back into the process of morphosyntactic linearisation. Of course, it is most likely that the grammatical system must integrate elements of both views compared here—i.e., phonological filtering and competition as well as interactive phonological feedback. We leave this as an open question for future work, and offer Tagalog adjective/noun word ordering as an example of a type of variable case study that could illuminate crucial interactions between morphophonological alternations and morphosyntactic order in natural language.

2.3 PHONOLOGICAL CONDITIONS ON WORD ORDER

The third interface issue is discussed at length in Shih & Zuraw 2017:§6.

REFERENCES


12 Or, in OT terms, that some phonological constraints can outrank syntactic ones in the assessment of word order output candidates (e.g., OCP » ALIGN (Adj, L); for alignment in OT syntax, see e.g., Woolford 2007).


BENOR, SARAH BUNIN; and ROGER LEVY. 2006. The Chicken or the Egg? A Probabilistic Analysis of English Binominals. Language 82.233–278.


EHRET, KATHERINA LUISA; CHRISTOPH WOLK; and BENEDIKT SZMRECŞÁNYI. 2014. Quirky quadratures: on rhythm and weight as constraints on genitive variation in an unconventional dataset. English Language & Linguistics 18.263–303.


