FUNCTION WORDS: IMPLICATIONS FOR THE SYNTAX-PHONOLOGY INTERFACE

Heather Newell (UQAM) & Tobias Scheer (UCA)

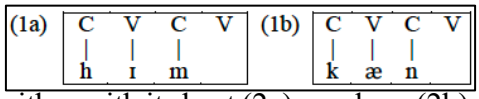
PUNCH LINE Phonological variation between weak and strong forms of function words (*her/'ər*, *him/'əm*) is due to a combination of phonological underspecification of functional items in the lexicon and the grammaticalization of syllabic space. Stress assignment is realized as the addition of syllabic space (depending on the theory: x-slots, moras, CV units) to phonological strings. Addition of this space to cycles containing phonologically underspecified functional morphemes causes their pronunciation as independent full forms. In its absence the reduced form is produced. This allows for a theory where variation between weak and full pronunciations of function words is directly linked to their morpho-syntactic structure/position. A further advantage of such an analysis is that it is fully modular; Phonology does not need to distinguish between so-called functional and lexical items.

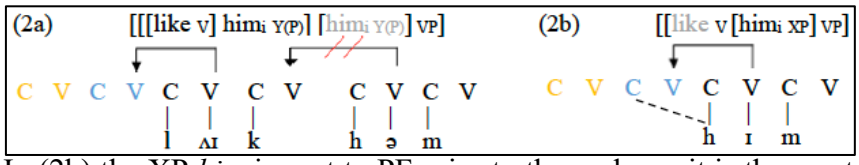
BACKGROUND *Function Words:* The current standard view of how function words (pron., prep., art. etc.) are treated by the phonology-syntax interface is that they are disregarded. Levels of the prosodic hierarchy that represent the mapping between syntax and phonology (the Prosodic Word (PW) and above) are argued to not be projected around function words (lexical X^0s do, but functional X^0s do not project PWs : Selkirk 1996, 2011 (MATCH), Truckenbrodt 2007 (WRAP)). But, recent work (from within OT) has successfully argued that the interface (spell-out) cannot be responsible for the distinction between functional and lexical phonology; the difference must be lexical in kind. Tyler (2019) proposes that weak functional items are lexicalized with prosodic subcategorization frames (see also Inkelas and Zec 1993). We advance this theory of a lexical basis for the distinction between lexical and functional words, but offer a more fine-grained and explanatory account for the pronunciation of weak vs full functional items. *Syllabic space:* Autosegmental phonological analyses appeal to the independent lexicalization of segments and syllabic space (Goldsmith 1976). Independently stored syllabic space underlies analyses in both Rule-Based and OT phonology to account for (e.g.) Arabic binyanim (since McCarthy 1979, 1981) or reduplication (Marantz 1982). Syllabic space is also shown to be the exponent of word stress (e.g. Chierchia 1986, Larsen 1998, Szigetvári & Scheer 2005, Bucci 2013, Ségéral & Scheer 2008), and the left edge of words (Lowenstamm 1999, Scheer 2014, d'Alessandro & Scheer 2015). The latter grammaticalized autosegmental elements are crucial to the analysis presented here.

OUTSTANDING PROBLEMS Any theory of the lexical nature of weak vs. strong function words must account for both the phonological and morpho-syntactic properties of these items. Among the important generalizations missed if one assumes that left/right subcategorization frames are responsible for the phonological output of weak function words (as in Tyler 2019) are: (I) Weak function words may cliticize to the left or right depending on their morphosyntactic environment (e.g. clitic pronouns in French (*Prends-**le*** 'Take **it**' vs. *Tu **le**-prends* 'You take **it**'). Clearly, the prosodic weakness of many clitics emerges regardless of their linear position. (II) Function words (like many items dubbed lexical words) are often demonstrably not X^0s, but are rather XPs (Cardinaletti & Starke 1999, Kayne 2010, Leu 2014). This fatally undermines theories like MATCH and WRAP which propose that the weakness of function words derives from a sensitivity at the phono-syntactic interface to the distinction between lexical and functional X^0s. (III) Functional items are variably pronounced as full or weak. A subcategorization frame account of this type of alternation is only viable if weak functional items are allomorphs of their full forms, i.e. if they have different underlying lexical specifications. In THE PROPOSAL section below, we demonstrate that some weak/strong alternations are fully phonological. Phonological alternations are indicative of *a single lexical underlier* for these items. The clitic nature of weak forms must therefore emerge from phonological properties that are shared with their full forms; they are not allomorphic alternations. *N.B.* Function words often display allomorphy (this is linked to the fact that function words, even though they are XPs, never evidence cyclic phonological domains (lexical words do: ex. [[[[*govern*] *ment*] *less*] *ness*]), even when overtly multi-morphemic (e.g. German [je-d-er]), except in the clear compound-like cases, e.g. *themselves*). Alternations such as *she~her* are clearly not derivable in the phonology. This variation therefore tells us little about the relationship between the phonology-syntax interface and the clitic nature of function words. Consequently, this presentation focuses on clearly

phonological variation, as it allows us to separate out behaviour specific to purely phonological processes that occur after morphology/vocabulary-insertion is complete.

THE PROPOSAL (A) Morphemes in the lexicon may be underspecified for the syllable structure needed to pronounce their segmental items. The insertion of CV-space that is the exponent of grammatical information explains certain phonological alternations. In English, for example, regular phonological processes are at play in full~clitic alternations. [h] and aspirated voiceless stops ([C^{h}]) are only pronounced word-initially or in the onset of a stressed syllable (Zwicky 1970, Kaisse 1985) ([h]*ilárious, ve*[h]*ícular* (cf. vé~~h~~icle), [t^{h}]*ell, a*[t^{h}]*ómic* (cf. á[t]om)). In CVCV phonology, this is due to the fact that [h]/[C^{h}] must be preceded by empty syllabic space; they are geminate-like (cf. Ségéral & Scheer 2008). Vowels are reduced in unstressed syllables (when they are short) and full when they are stressed (long). The variable pronunciation of (e.g.) *him* [hɪm]/*her* [həɹ] as [(ə)m̩]/[(ə)ɹ̩] or *can* [k^{h}æn]/*to* [t^{h}uw] as [kən]/[tə] tracks these regular phonological rules. (B) CV-space is inserted in all cases where stress is assigned to a domain (incl. Focus). (C) The left-edge of independently interpreted phases/phonological cycles is endowed with an empty CV. A morpheme that does not contain enough syllabic space to enable its full pronunciation (e.g. (1)) may undergo vocabulary insertion either with its host (2a), or alone (2b), depending on its morpho-syntactic position. A clitic (object) pronoun (2a) is syntactically smaller than a full pronoun and raises into a position in which it is interpreted in the domain of its host (e.g. Cardinaletti & Starke 1999, Déchaine & Wiltchko 2002, Preminger 2009). A full pronoun (2b) is morpho-syntactically larger, leading to semantic and distributional distinctions. Importantly, this entails that the pronoun in (2b) is an independent cycle for phonological interpretation.





In (2b) the XP *him* is sent to PF prior to the verb, as it is the most embedded cyclic domain. As such, it will be assigned a left-edge CV at the PF interface. The English stress algorithm will then apply (as *him* has satisfied *word minimality*), inserting a CV to the left of the stressed vowel. As [h] is preceded by an empty vocalic position, it is licensed and pronounced. Compare this to (2a), where the verb is interpreted prior to the clitic (as it is the most embedded cycle). It will therefore be assigned a left-edge CV and will be stressed (CV). In a second cycle, *him* is interpreted. As it is not leftmost in its domain at PF (it is linearized in relation to content it c-commands (Newell 2008)) no empty CV intercedes between the verb and the clitic. As such, [h] will remain unpronounced, and *him* will not affect the stress position of the verb (cf. Level 2 affixes). Vowels will be stressed only if they have access to the requisite syllabic space. This account will be shown to extend to a wide range of clitic/weak forms; to the C/C^{h} distribution of initial Cs in Eng., to variable hiatus resolution in French (*l*[ə] *garçon/l'arbre* vs *r*[ə]*hausser/r-ajouter*), and to pronominal/determiner reduction/lenition in other languages (e.g. Haitian Creole *mwen/m*, Swedish *den/ren*), underlining the phonological nature of these alternations; Non-allomorphic variation in function word pronunciation is derivable via regular phonological rules in each of these languages. It also captures, via a confluence of morpho-syntactic and phonological means, *why* certain function words are clitics, and *when* they will emerge as full-forms.

THE IMPORTANCE OF THIS WORK It is only possible to have an explanatory account of function word phonology once we consider the full implications of morpho-syntactic accounts of functional structure in combination with independently supported (non-ad-hoc) representational theories of phonology. The account proposed here is modular, relies only on regular, independently motivated operations specific to spell-out and phonology and has implications for the study of cross-linguistic function-word phonology. This analysis also operates as a window into where we expect to find fine-grained phonological evidence of cyclic spell-out domains.