

How Many Levels of Phrasing? Empirical Questions and Typological Implications

Sam Hellmuth
University of York

1. Introduction

This paper explores whether a phase-based approach to the syntax-phonology interface can account for the prosodic phrasing facts of Egyptian Arabic (EA), and the potential implications of such an analysis for our understanding of the nature of the prosodic hierarchy (in EA if not beyond). Specifically, we examine the phrasing facts observed in a corpus of syntactically complex and prosodically heavy VOO read speech sentences, in comparison to previously published generalisations for parallel SVO sentences, for evidence to disambiguate between the predictions of competing conceptions of the syntax-phonology interface: ‘edge-based’, in which the phonology is sensitive to the edges of lexical maximal projections (Selkirk 1986) vs. ‘phase-based’, in which prosodic structure is mapped from functional projections (Chomsky 2001). The evidence proves to be finely balanced, but we do find that the mapping must, under either analysis, result in a level of phonological representation mediating between syntactic structure and phonetic implementation, to account for observed prosodic minimality effects. In addition, the detail of the phonetic implementation suggests that the intervening prosodic representation must be one which allows recursive structure. We conclude by suggesting that EA prosodic structure may consist of fewer levels of phrasing than proposed even in recent, universally restricted, conceptions of the Prosodic Hierarchy (Itô & Mester to appear).

2. Background to the study

Egyptian Arabic (EA) is defined here as the dialect of Arabic spoken in Cairo, which functions as a national standard throughout Egypt (Haeri 1996, Bassiouney 2009). The segmental and metrical phonology of EA are well-described (Watson 2002). Intonationally, EA is a stress accent language in which the function of pitch is purely post-lexical, but differs from many intonation languages in having an unusually rich default accent distribution pattern (Chahal & Hellmuth to appear). Prior work on EA, within the Autosegmental-Metrical framework of intonation (Gussenhoven 2004, Ladd 2008), has shown that EA displays a pitch movement (analysed as a tonal target, or pitch accent) on almost every content word, across a range of speaking styles and contexts (Hellmuth 2006). This distribution can be formalised as obligatory association of a phonological target (a default LH* pitch accent) with the head foot of every Prosodic Word (PWd) (Hellmuth 2007). Syntactically, although it displays both SVO and VSO surface word orders (Brustad 2000), EA can be analysed as having basic SVO word order (Benmamoun 2000, Edwards 2009).

The prosodic phrasing facts of EA have been previously investigated, both instrumentally (Norlin 1989) and in terms of the syntax-phonology interface (Hellmuth 2004, 2007, to appear). A typical EA declarative utterance shows declination in the relative height of the pitch peaks which are realised on each content word, with the last accent in the utterance often realised somewhat lower than the declination trend would predict, in the effect known as final lowering (Liberman & Pierrehumbert 1984). The most consistent cue to prosodic phrasing in EA is suspension of this global declination trend, in the form of upstep of a phrase-final accent and/or local pitch reset after a phrase-boundary, frequently though not obligatorily accompanied by pre-boundary lengthening (Hellmuth to appear). Hellmuth (2004) investigated phrasing in a corpus of read speech SVO sentences in which syntactic complexity and prosodic weight was systematically varied (following Frota *et al.* 2007) and found that

3. Methodology

Eight VOO (ditransitive) sentences were constructed in which the direct object and indirect object were both syntactically complex and prosodically heavy (see Table 1). The direct object was the same in all eight sentences, but the indirect object varied, as shown in Table 2. The potential boundary positions of interest are the ‘after DirObj’, ‘after Loc(ative)’ and ‘within IndObj’ positions. The VOO dataset also includes locative PPs which are open to more than one syntactic analysis: in (4) the sentence-medial locative PP [fi 'gamʕat ju'na:n] ‘in university-Greece’ is analysed as a vP-adjunct (cf. Adger 2003), with the indirect object [li tʕalaba 'nubaha min 'ru:ma] ‘to students clever from Rome’ extraposed by rightward movement; in (5) the sentence-medial locative PP is analysed as a VP-adjunct (cf. Adger & Tsoulas 2000). The edge-based phrasing predictions for the two syntactic structures are the same: (V DirO) (AdjP) (IndO) (PP). A canonical phase-based prediction for the structure in (4) is (V DirO) (AdjP) (IndO PP); whereas for (5) it is (V DirO AdjP IndO PP).

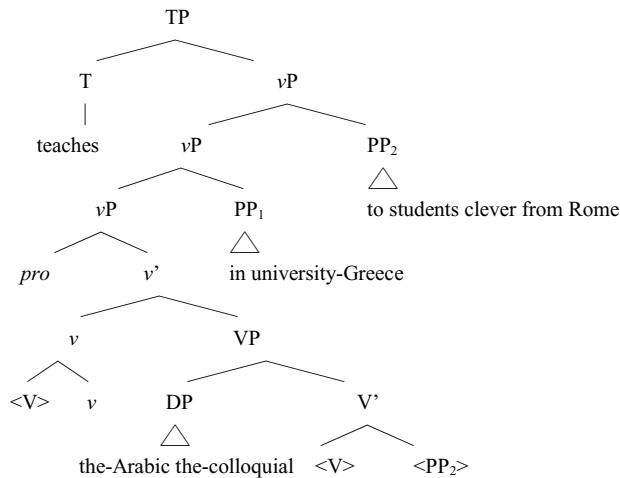
Table 1: Structural template of VOO target sentences (showing potential boundary positions)

| verb | direct object | | indirect object | |
|---------------------------------|------------------------------------|---------------------------------------|---|--|
| bijʕallim <i>teaches.3ms</i> | il-ʕara'bijja <i>the-arabic</i> | il-ʕam'mijja <i>the-colloquial</i> | fi 'gamʕat ju'na:n <i>in university-Greece</i> | li tʕalaba 'nubaha min 'ru:ma <i>to students clever from Rome</i> |
| | | after DirObj ↑ | after Loc ↑ | within IndObj ↑ |

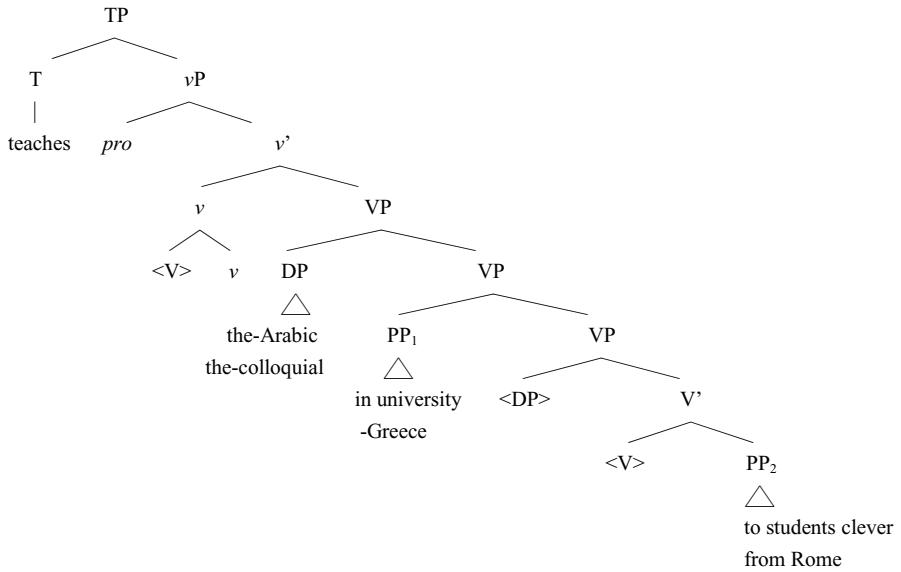
Table 2: Internal structure of the indirect object in the 8 target sentences analysed in the present study

| | | | | | | |
|-----|--------|--------------|--|--|------------------|---------------------------------|
| 038 | li | tʕalaba | 'nubaha | min | 'ru:ma | |
| | [to [| [students | [clever] _{AP}] _{NP} | [from [Rome] _{DP}] _{PP} | |] _{DP}] _{PP} |
| 040 | li | li tʕul'la:b | 'nubaha | min | 'ru:ma | |
| | [to [| [students | [clever] _{AP}] _{NP} | [from [Rome] _{DP}] _{PP} | |] _{DP}] _{PP} |
| 042 | li | li tʕalaba | muʔaddi'bi:n | min | 'ru:ma | |
| | [to [| [students | [polite] _{AP}] _{NP} | [from [Rome] _{DP}] _{PP} | |] _{DP}] _{PP} |
| 044 | li | li tʕul'la:b | muʔaddi'bi:n | min | hu'landa | |
| | [to [| [students | [polite] _{AP}] _{NP} | [from [Holland] _{DP}] _{PP} | |] _{DP}] _{PP} |
| 052 | li | li tʕalaba | 'nubaha | min | 'gamʕat hu'landa | |
| | [to [| [students | [clever] _{AP}] _{NP} | [from [university-Holland] _{DP}] _{PP} | |] _{DP}] _{PP} |
| 048 | li | li tʕalaba | muʔaddi'bi:n | min | 'gamʕat 'ru:ma | |
| | [to [| [students | [polite] _{AP}] _{NP} | [from [university-Rome] _{DP}] _{PP} | |] _{DP}] _{PP} |

(4)



(5)



The 8 sentences were pseudo-randomised and interspersed with distractor sentences from other experiments. They were presented in Arabic typescript, using EA spelling conventions to elicit colloquial productions (Siemund *et al.* 2002). Recordings were made in Cairo with 6 female speakers of EA who were all mother tongue speakers of EA, born and raised in Cairo. They were aged between 21-34 years and none had any speech or hearing difficulties. The recordings were made with an AKG headset condenser microphone, in digital .wav format at 44.1KHz 16bit, re-sampled at 22.05KHz 16bit. Each sentence was produced 3 times by each of the 6 speakers, yielding 108 tokens for qualitative and quantitative analysis (disfluent tokens were not excluded) using Praat (Boersma & Weenink 2009).

Information was hand-labelled by the author on three tiers, based on auditory impression and inspection of the pitch trace and spectrogram, as illustrated in Figure 2 below. The utterance was first segmented into PWds, with enclitics grouped with a preceding host content word. A Praat script was used to automatically identify the pitch maximum within each labelled PWd and a peak label inserted at that point. A further Praat script extracted the f0 value in Hertz (measured in a hand-corrected Pitch object) at each labelled peak position. Note that the f0 peak identified for a PWd at a boundary could reflect either a pitch accent on the word itself or a phrase tone at its right edge (e.g. if it bears a high phrase tone). Finally, qualitative Break Indices and cue labels³ were hand-labelled on two more tiers (cf. Hellmuth to appear).

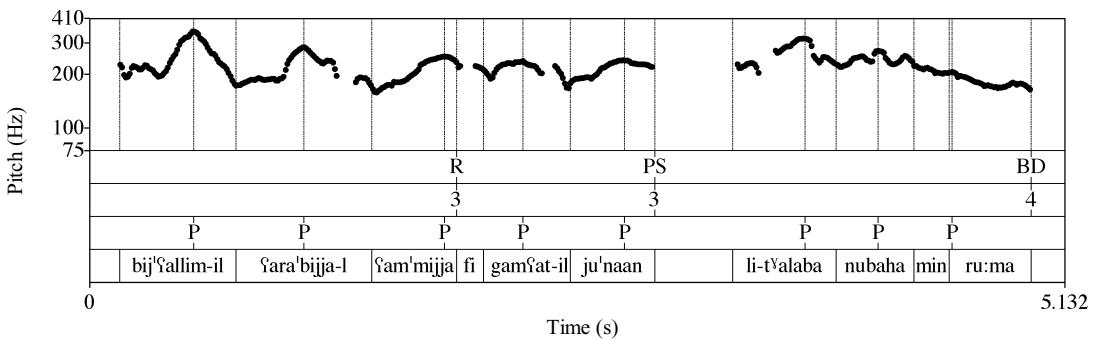


Figure 2: Sample labelled TextGrid of ditransitive target sentence (token 038fhg1).

³ R= following reset; P = pause; S= suspension of downstep; B=boundary tone; D= final lowering.

4. Results

Phrasing generalisations, based on the qualitative Break Indices labelling, are presented first, followed by details of the relative prosodic strength of boundaries observed in different structural positions, based on quantitative peak height measurements. The most common phrasing pattern, observed in 65% of tokens, shows a boundary after the Direct Object and after the sentence-medial locative, but not within the Indirect Object as in (6) below⁴. The second most common pattern, observed in 23% of tokens, adds an additional boundary within the Indirect Object as in (7). The remaining 12% of tokens (N=13) show either unusual phrasings due to hesitation or an unexpected pattern (e.g. a boundary after the verb) and are not analysed further here.

- (6) (V DirO) (AdjP) (IndO PP) observed in 65% of tokens
 (7) (V DirO) (AdjP) (IndO) (PP) observed in 23% of tokens

A breakdown of the observed phrasing choices by sentence type (with a short vs. long final PP), illustrated in Figure 3 below, reveals that the number of PWds in the final PP is a factor in determining the phrasing pattern of the sentence. Phrasings in which a short final PP is phrased as an independent prosodic phrase, in violation of BINMAP, are avoided, and though they occur they are rare⁵. A breakdown of phrasing choices by speaker shows that two speakers, fhg and fhx, are somewhat consistent in using only the phrasing pattern in (6), whereas the others show more variation in their choices. The speakers are presented in order of increasing speech rate, showing that speech rate does not determine this variation.

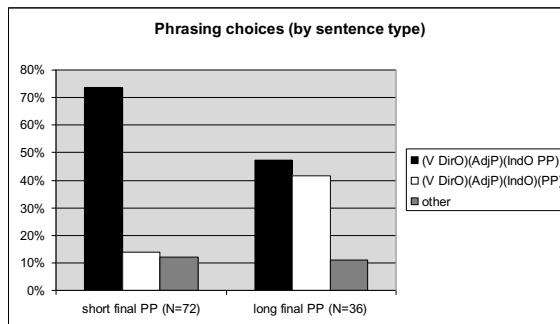


Figure 3: Phrasing choices by sentence type.

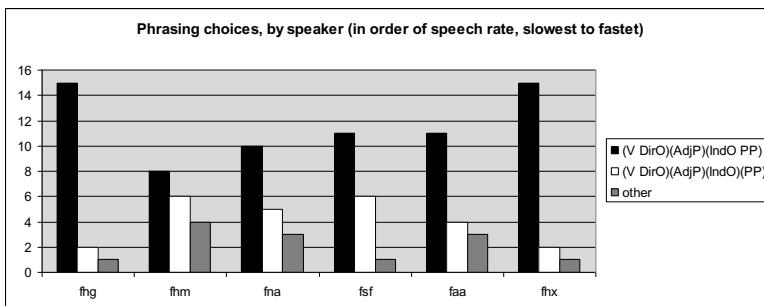


Figure 4: Phrasing choices by speaker (presented in order of increasing speech rate)⁶.

⁴ The token illustrated in Figure 2 is an example of this phrasing pattern.

⁵ There are 10 tokens in which a short final PP is realised in its own prosodic phrase, in violation of BinMaP. They are produced by four of the six speakers: faa, fhm, fna, fsf. There is no obvious correlation with speech rate, though we note that these are the four speakers who show a more mixed range of phrasing choices overall, as shown in Figure 4.

⁶ As calculated for the SVO dataset analysed in Hellmuth (to appear), which was recorded in the same session.

Assuming the syntactic structure in (4), the observed generalisations support either a phase-based analysis, which predicts boundaries in after-DirObj and after-Loc position only as in (6), or an edge-based analysis with prosodic minimality (BINMAP) preventing formation of short final prosodic phrases.

Analysis of the relative height of peaks throughout each utterance, however, broken down by speaker, reveals subtle differences among speakers in how the phrasing patterns are realised. Figures 5 and 6 show mean peak heights, by speaker, for sentences with a short and long final PP, respectively. Four speakers (fhg, fhm, fhx and fsf, marked with solid lines) show a pattern in which the degree of pitch reset (whether upstep before the boundary or reset after it) is greater in after-Loc position (between peaks 5 and 6) than in after-DirObj position (between peaks 3 and 4). The other two speakers (faa and fna, marked with dotted lines) show the reverse pattern; for them the later (after-DirObj) boundary is marked by relatively weaker cues, with only a slight reset in each of the two boundary positions.

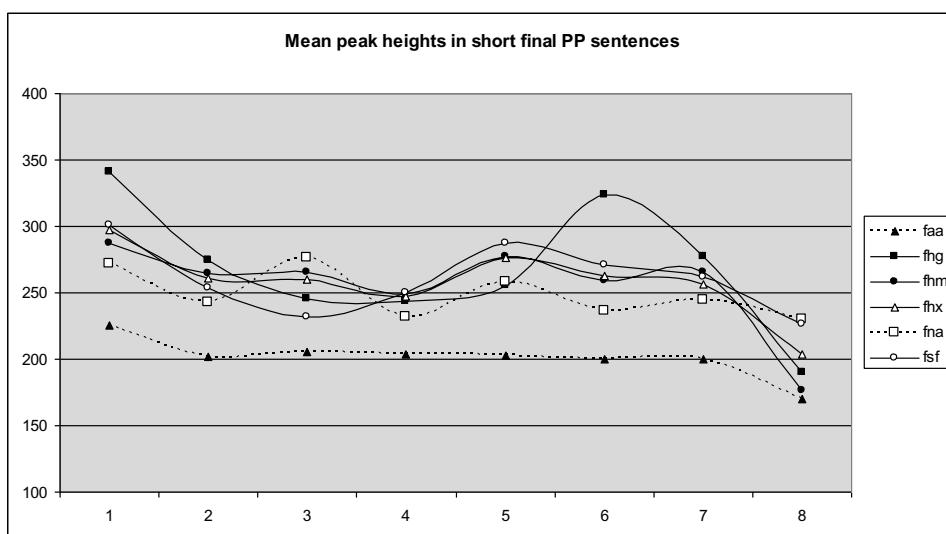


Figure 5: Mean peak heights in sentences with a short final PP, by speaker.

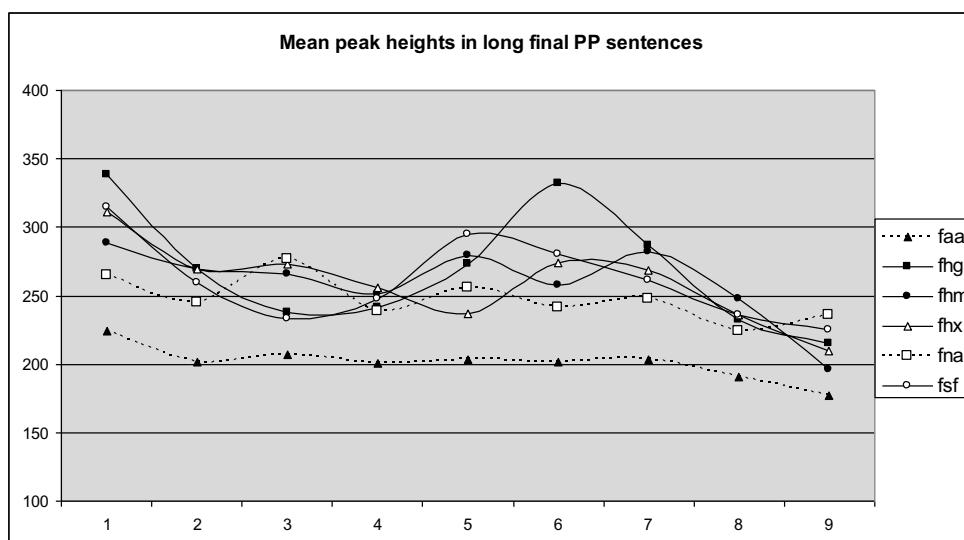


Figure 6: Mean peak heights in sentences with a long final PP, by speaker.

This difference in the phonetic realisation of boundaries suggests that all boundaries are not equal. One solution would be to propose that the stronger vs. weaker cues indicate boundaries at different levels of the Prosodic Hierarchy. This solution is rejected in Hellmuth (to appear) on empirical grounds, in favour of an analysis in terms of prosodic subordination between adjacent phonetic implementation domains (Truckenbrodt 2007), which we also adopt here.

The new VOO dataset sheds further light on the observed relations among prosodic domains, however. In the SVO dataset (in Hellmuth to appear) it was always the later boundary which was more deeply embedded, and showed weaker prosodic cues. This pattern could be explained *either* in terms of a purely linear relationship among successive prosodic domains in a flat (strictly layered) structure, or in terms of hierarchical relations among recursive prosodic domains which reflect the degree of syntactic embedding (cf. Ishihara 2007 fn. 12). The pattern observed in the present VOO dataset, in which the earlier (more embedded) boundary is the stronger of the two, is not predicted in a conception of the Prosodic Hierarchy that maintains the Strict Layer Hypothesis (Selkirk 1984). It is predicted in a phase-based analysis in which spelled-out material is mapped to prosodic structure in successive cycles, reflecting the relative embedding of syntactic constituents, or in an XP-based analysis which permits recursion (Truckenbrodt 1999, Selkirk to appear).

The reverse boundary strength pattern observed in the phonetic realisation patterns of the two remaining speakers (faa and fna) presents a puzzle. The possibility of alternative syntactic analyses of the VOO target sentences, discussed in section 3 above, rests on differing treatment of the sentence-medial locative PP, as either a vP- or VP-adjunct. Since these differing treatments can be said to give rise to different interpretations, it may be that both are available to speakers. If so, then the four speakers whose realisations show a stronger boundary in after-Loc position are treating the locative as a vP-adjunct, as in (4) above, whereas the two speakers whose realisations show a stronger boundary in after-DirObj position are treating the locative as a VP-adjunct, as in (5). A problem for the phase-based analysis, however, is that the two ‘VP-adjunct’ speakers do nonetheless realise a boundary in both XP-edge positions, with the earlier boundary stronger than the later. This phrasing is not predicted in a phase-based analysis, if their realisations map from the syntactic representation in (5), but is predicted in an edge-based analysis. In future work we hope to test perceptually whether the phrasing realisations of the two sets of speakers give rise to differing interpretations, in order to further disambiguate between the two theoretical approaches.

5. Conclusion

The phrasing generalisations observed in the present dataset suggest that it is possible to analyse the phrasing facts of EA in either an edge-based or a phase-based analysis, though the edge-based mapping is able to account for a wider range of the possible phrasings observed (that is, both 6) and 7)). In either case we find empirical evidence that the result of the syntax-phonology mapping must be some form of phonological representation, rather than direct spellout to phonetic implementation, since the resulting prosodic structure shows sensitivity to prosodic minimality constraints: phrasings of a single-word final PP as an independent prosodic phrase are strongly resisted in the current dataset. This presents independent motivation for retention of a Prosodic Hierarchy of some form (cf. Scheer 2009). The variable cues to phrasing observed in EA allow us to disambiguate between a strictly layered vs. recursive conception of the Prosodic Hierarchy, since the phonetic detail of relative prosodic boundary strength is shown here to reflect the degree of syntactic embedding, and not the purely linear order in which boundaries are realised. Finally, although the realisations of two of the speakers present a potential challenge to a phase-based approach, they also offer up the future possibility of a means of disambiguating between the two competing theories.

All of the patterns observed in the present dataset can be explained in a phase-based analysis which assumes spellout of vP and CP as prosodic phrases but excludes adjuncts from the spellout domain (Ishihara 2007), assuming the syntactic structure in (4), and allowing for recursive prosodic structure. What is spelled out in the theory of Ishihara (2007), however, is a single level of prosodic phrasing (which he calls MaP), regardless of whether it is triggered by a vP or CP phase. This contrasts with Match Theory (Selkirk to appear) which argues for distinct mappings from distinct components of syntactic structure (cf. Selkirk 2005): functional projections (CP and ForceP) map to Intonational

Phrases (MatchClause), whereas lexical projections (XPs) map to Phonological Phrases, equivalent to MaP (MatchPhrase). Although the prosodic minimality effects observed in the current dataset confirm that an intervening prosodic representation is indeed required, it seems that the EA facts can be analysed adequately by appeal to either lexical (edge-based) or functional (phase-based) projections, but not both. Although some languages show distinct cues to phrasing at different levels of the prosodic hierarchy, this is not the case in EA, as outlined in section 2 above. The prominence hierarchy proposed for English, of lexical accent vs. pitch accent vs. nuclear accent (Beckman & Pierrehumbert 1986), marking accentual heads of PWds, MaPs and IPs respectively, runs into immediate problems in EA, since the first two levels of prominence, lexical accent vs. pitch accent are in EA routinely conflated (Chahal & Hellmuth to appear). Since we are not aware of any instrumental evidence confirming the presence of phonetic marking of nuclear prominence other than under focus conditions⁷, we suggest that spellout of a single level of phrasing may be sufficient for analysis of the phrasing facts of EA. This would yield a ‘prosodic hierarchy’ composed only of PWds (mapped from morphosyntactic words) and MaPs (mapped from spellout domains). In turn, phenomena such as boundary tone combinations may in EA prove to be better analysed as markers of turns (mapped from discourse structure) than of a distinct IP level domain.

Acknowledgements

The author wishes to thank the Egyptian Arabic speakers for their time and patience, and members of the audience at WCCFL 28 and ZAS Berlin, and Bernadette Plunkett and George Tsoulas for comments.

References

- Adger, David. 2003. *Core Syntax*. Oxford: Oxford University Press.
- Adger, David & George Tsoulas. 2000. Aspect and lower VP adverbials. In Alexiadou, Artemis & Peter Svenonius (eds.) *Adverbs and Adjunction*. 1-18. Potsdam: Linguistics in Potsdam 6.
- Bassiouney, Reem. 2009. *Arabic sociolinguistics*. Edinburgh: Edinburgh University Press.
- Beckman, Mary & Janet Pierrehumbert. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3 255-309.
- Benmamoun, Elabbas. 2000. *The feature structure of functional categories*. Oxford: OUP.
- Brustad, Kristen. 2000. *The syntax of spoken Arabic: a comparative study of Moroccan, Egyptian, Syrian and Kuwaiti dialects*. Washington: Georgetown University Press.
- Chahal, Dana & Sam Hellmuth. to appear. Comparing the Intonational Phonology of Lebanese and Egyptian Arabic. *Prosodic Typology. Volume 2*. Oxford: Oxford University Press.
- Chomsky, Noam. 2001. On derivation by phase. In Kenstowicz, M. (ed.) *Ken Hale: a life in language*. 1-52. Cambridge, Mass.: MIT Press.
- Edwards, Malcolm. 2009. Word order in Egyptian Arabic: form and function. In Owens, Jonathan & Alaa Elgibali (eds.) *Information Structure in Spoken Arabic* Oxford: Routledge.
- Fox, Danny & David Pesetsky. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics*. 31 1-46.
- Frota, Sonia, Mariapaola D'Imperio, Gorka Elordieta, Pilar Prieto, & Marina Vigario. 2007. The phonetics and phonology of intonational phrasing in Romance. *Amsterdam Studies in the Theory and History of Linguistic Science IV*. 282 131-154.
- Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: CUP.
- Haeri, Nilofaar. 1996. *The sociolinguistic market of Cairo: gender, class and education*. London: Keegan Paul International.
- Hellmuth, Sam. 2004. Prosodic weight and phonological phrasing in Cairene Arabic. *Proceedings of Annual Meeting of Chicago Linguistic Society*. 40.
- Hellmuth, Sam. 2006. *Intonational pitch accent distribution in Egyptian Arabic*. Unpublished PhD thesis, SOAS.
- Hellmuth, Sam. 2007. The relationship between prosodic structure and pitch accent distribution: evidence from Egyptian Arabic. *The Linguistic Review*. 24(2). 289-314.
- Hellmuth, Sam. to appear. Variable cues to phrasing: finding edges in Egyptian Arabic. In Borowsky, Toni, Shigeto Kawahara, & Mariko Sugahara (eds.) *Prosody matters: essays in honor of Lisa Selkirk*. London: Equinox.

⁷ Following Ishihara (2007), focus-induced prominence may be due to a separate mechanism than that which gives rise to prosodic phrasing.

- Ishihara, Shinichiro. 2007. Major phrase, focus intonation and multiple spellout. *The Linguistic Review*. 24(2).
- Itô, Junko & Ralf-Armin Mester. to appear. Trimming the prosodic hierarchy. In Borowsky, Toni, Shigeto Kawahara, Tahahito Shinya, & Mariko Sugahara (eds.) *Prosody matters: Essays in honor of Elisabeth Selkirk* London: Equinox.
- Ladd, D. Robert. 2008. *Intonational Phonology*. Cambridge: Cambridge University Press.
- Liberman, Mark & Janet Pierrehumbert. 1984. Intonational variance under changes in pitch range and length. In Aronoff, M. & R. T. Oehrle (eds.) *Language Sound Structure: Studies in Phonology present to Morris Halle*. 157-233. Cambridge, M.A.: MIT Press.
- Norlin, Kjell. 1989. A preliminary description of Cairo Arabic intonation of statements and questions. *Speech Transmission Quarterly Progress and Status Report*. 1 47-49.
- Scheer, Tobias. 2009. Intermodular Argumentation and the Word-Spell-Out-Mystery. In Grohmann, Kleanthes (ed.) *Explorations of Phase Theory: Interpretation at the Interfaces*. 23-65. Berlin: Mouton de Gruyter.
- Selkirk, Elisabeth O. to appear. The Syntax-Phonology Interface. In Goldsmith, John, Jason Riggle, & Alan Yu (eds.) *The Handbook of Phonological Theory*.
- Selkirk, Elisabeth O. 1984. *Phonology and syntax: the relation between sound and structure*. Cambridge, Mass.: MIT Press.
- Selkirk, Elisabeth O. 1986. On derived domains in sentence phonology. *Phonology Yearbook*. 3 371-405.
- Selkirk, Elisabeth O. 2000. The interaction of constraints on prosodic phrasing. In Horne, Merle (ed.) *Prosody: theory and experiment*. 231-262. Dordrecht: Kluwer.
- Selkirk, Elisabeth O. 2005. Comments on intonational phrasing in English. In Frota, Sonia, Marina Vigario, & Maria J. Freitas (eds.) *Prosodies. (Selected papers from the Phonetics and Phonology in Iberia Conference, 2003)*. *Phonetics and Phonology Series*. Mouton de Gruyter.
- Siemund, R, B Heuft, K Choukri, O Emam, E Maragoudakis, H Tropf, O Gedge, S Shammass, A Moreno, A Nogueiras Rodriguez, I Zitouni, & D Iskra. 2002. OrienTel - Arabic speech resources for the IT market. LREC 2002 (Arabic workshop).
- Truckenbrodt, Hubert. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry*. 30(2). 219-255.
- Truckenbrodt, Hubert. 2007. Upstep on edge tones and on nuclear accents. In Riad, Tomas & Carlos Gussenhoven (eds.) *Tones and Tunes: Experimental studies in word and sentence prosody. Volume 2: Phonology and Phonetics*. 349-386. Berlin: Mouton de Gruyter.
- Watson, Janet C. E. 2002. *The phonology and morphology of Arabic*. Oxford: OUP.

Proceedings of the 28th West Coast Conference on Formal Linguistics

edited by Mary Byram Washburn,
Katherine McKinney-Bock, Erika Varis,
Ann Sawyer, and Barbara Tomaszewicz

Cascadilla Proceedings Project Somerville, MA 2011

Copyright information

Proceedings of the 28th West Coast Conference on Formal Linguistics
© 2011 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-441-6 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Hellmuth, Sam. 2011. How Many Levels of Phrasing? Empirical Questions and Typological Implications.
In *Proceedings of the 28th West Coast Conference on Formal Linguistics*, ed. Mary Byram Washburn et al.,
258-266. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2458.