Coming to the party late
On the timing and limitations of word building

Paula Fenger
9th February 2021, SyntaxLab Cambridge

• Overarching question: how does functional syntactic material get mapped into different modules of grammar?

• The unit of interest: the word
  – Is word a coherent unit in grammar?

(1) “Ran inter a couple o’ mad trolls on the Polish border an’ I had a sligh’ disagreement with a 
vampire in a pub in Minsk, bu’ apart from tha’ couldn’t’a been smoother.”
(Harry Potter and the order of the Phoenix, chapter 20, p. 377)

• Ideal case: Syntax, Phonology, Morphology and Meaning all align ⇒ Word

(2) /vampire/
  a. Syntax: N 1 unit
  b. Phonology ( /ˈvæmpərə/ ) 1 unit
  c. Morphology: [vampire] 1 unit
  d. Meaning: 1 unit

• But, so many cases where one or the other does not align: mismatches, or words within words

(3) /couldn’t’a/
  a. Syntax: [TP T[PST] [NegP Neg [VP Vcan [VP V[PRF] [VP . . . ]]]]] 4 units
  b. Phonology ( /ˈkudntə/ ) 1 unit
  d. Meaning: t < s* (PST), ¬ (Neg), ◊ (can), t′ ⊆ t (PRF) 4? units

• Mismatches as basis, (3): The word is not a coherent unit of grammar in any component

• Mismatches as noise, (2) as basis: There is a coherent unit ⇒ Position defended here
• **Empirical domain**: Verbs + inflection: Rich variation in inflectional categories and (non)-mismatches

• The expression of *Tense, Aspect, Mood* *(TAM)*, and *Valency changing/Voice*:
  
  – Features in separate ‘words’, i.e. periphrastically, (4) for Dutch
  – Features in one ‘word’, i.e. affixal, (5) for Latin
  – Features in one-and-a-half ‘word’, i.e. *a word within a word*, (6) for Turkish

(4) Zonder Adriaan zou Bassie de schat niet **hebben kunnen vinden**  
Without Adriaan would Bassie the treasure not have could find  
‘Without Adrian Bassie could not have found the treasure’

(5) [**laud** -a -u -i -s -s -e -mus ]  
praise -TV -PRF -TV -PST -SUBJ -TV -1.PL  
‘Would we have been praised’  
(Embick and Halle 2004)

(6) [**ver** -il -ebil -ir ] -di ]  
give -PASS -ABIL -HAB -PST  
‘it could be given to him/her’

• **Proposal**: Word-building processes can be timed at different stages in the derivation
  
  – There is a coherent unit, it is built in the syntax, i.e. a *syntactic word*
  – Mismatches come from timing differences in word-building (before/after syntax, in phonology)
  – Syntactic word-building can blocked by phases

• Roadmap
  
  – Mismatches between syntax & phonology (focusing on Turkish, Sec. 1)
    o How do we know that Turkish has one-and-a-half words, but Dutch has separate words?
    o Units that are smaller than a ‘word’ align with syntactic constituents (syntactic words)
  – Look at detail at head movement (Sec. 2, 3)
    o Timing of operations: before or after spell-out
    o Empirical domain: cross-linguistic variation in ellipsis
  – Limitations of movement (Sec. 4)
    o Exploration of the relation with phases
    o Looking at correlation with phrasal phenomena, and cross-linguistic variation

1 **Arriving (very) Late: Word Building during and after syntax**

• How do we know if the correct analysis of (7) is (7a) or (7b)?

(7) V X₁ X₂:  
a. [ V - X₁ ] [ X₂ ]  
b. [ V - X₁ - X₂ ]

• There is a puzzle in Turkish with regard to phonological properties
  
  – **CAUS-PST**: Stress falls at the ‘end’ of the word, *vowel harmony* on all vowels, (8);
  – **PROG-PST**: Stress fall in the ‘middle’, *vowel harmony* includes **PST** in V, (9);
Coming to the party late 3 Paula Fenger

(8) \((\text{koş -} \text{tur} \text{-} \text{du})_{\omega}\) \(\text{run -CAUS -PST}\)

‘\(x\) makes \(y\) run’

(9) \((\text{kal -} \text{iyor} \text{-} \text{du})_{\omega}\)

stay -PROG -PST

‘was staying’

• I show that this variation is not random
  – The smaller phonological constituent (w.r.t. stress) matches up with a syntactic constituent
  – features matter. There is a cut-off point after aspect

• I derive (in Sec. 1.3) the variation by assuming that: (i) word-building can be timed at different stages in the derivation, and (ii) is sensitive to phases, which together lead to variation in wordhood

1.1 Wordhood in Phonology

• Generally, Turkish stress is assigned to the last syllable of a ‘word’ (Lees 1961, Lewis 1967, Sezer 1983).

(a) \(\text{kit\'ap}\)

‘book’

(b) \(\text{kitap\'l\'ik}\)

‘bookcase’

(c) \(\text{kitab\'l\'ar}\)

‘bookcases’

(d) \(\text{kitab\'lar\'im}\)

‘my bookcases’

(e) \(\text{kitab\'lar\'im\'iz}\)

‘our bookcases’

(f) \(\text{kitab\'lar\'im\'iz\'dan}\)

‘from our bookcases’

• This can be informally represented in the following rule (based on Kabak and Vogel 2001, Güneş 2009).

(11) Apply stress to the last syllable of a word

• This is also the case for verbs with an affix.\(^1\)

  – There are no exceptions, all single TMA markers can receive stress:

(12) \((\text{bitir} \text{-} \text{e\'cek}) [\text{FUT}]_{\omega}\) \(\text{finish -FUT}\)

‘s/he will finish’ Kornfilt

(13) \((\text{kal -} \text{du})_{\omega}\)

stay -PST

(14) \((\text{koş -} \text{tur})_{\omega}\)

run -CAUS

‘s/he stayed’

(1997)

‘x makes \(y\) run’

• It is also possible to combine certain morphemes, in which case stress shifts, (8):

  – Stress makes reference to a word boundary, not a specific morpheme: cf. CAUS in (14) vs. (8).

• Importantly, there is another phonological process that picks out ‘words’: vowel harmony (VH)\(^2\)

\(^{1}\text{Data from sources cited and consulting G. Güneş and H. Sevgi.}\)

\(^{2}\text{I do not intend to give a complete overview of vowel harmony patterns in Turkish, but what is important here is when vowel harmony can be used to detect a single unit. Having multiple vowel harmony domains is not necessarily indicative of a single unit, since there are various vowels that are specified for vowel quality and as such can start a new vowel harmony domain inside the same word (Clements and Sezer 1982). However, when there is a single domain, it must mean that all vowels are in the same domain, i.e. in the same phonological word.}\)
– assimilatory process in which vowels of a word have match for certain features ([+/-back], [+/-round]) in Turkish
– When there is a single VH domain, there is a single ‘word’: cf. PST in (13) vs. (14)
– This is important, because these processes can pick out different units.

(15) **Turkish phonological processes (to be revised):**
   a. Apply stress to the last syllable of a word
   b. Apply vowel harmony to a word

• Certain TMA combinations lead to auxiliaries.
  – Affixes cannot be stacked with stress on the final syllable;
  – An auxiliary is used, and stress is on the main verb;

(16) a. *(yaz -iyor -a’çak)*ω
    write -PROG -FUT

[153x476] b. *(yaz -iyor)*ω *(ol -acak)*ω
    write -PROG AUX -FUT

‘s/he will be writing’

• Interestingly, in some cases there is no auxiliary present, (17a):
  – Stress pattern is the same as with the auxiliary cases in (16b)
  – But, PST harmonizes with element of the main verb VH domain, like in (17b)

(17) a. *(kal -iyor )ω -du)*ω
    stay -PROG -PST

[145x371] b. *(koş -tur -du)*ω
    run -CAUS -PST

‘was staying’

– Stress does not align: Stress falls in ‘the middle’ of the word, VH picks out the whole word.
  – This has been dealt with in phonology as an exceptional stress pattern (van der Hulst and van de Weijer 1991, Inkelas and Orgun 1998, 2003, Inkelas 1999, Kabak and Vogel 2001)
    ○ However, stress is not a property of a morpheme, but of the environment
  – It is a side effect of mapping to PF (building on Kornfilt 1997, Kahemuyipour 2006, Kamali 2011)
    ○ There are in fact multiple syntactic units in the mismatched ‘words’

⇒ This predicts that there should be syntactic evidence for multiple units.

### 1.2 Wordhood in syntax

• How to find syntactic words? Constituency tests (coordination, particle placement, …)
  – Units picked for constituency tests align with stress, not with VH ⇒ VH masks syntactic output

• Particles can occur in between different ‘words’

(18) Kitap { bile / da / mi } oku -yaca˘g -ım
    book even also Q read -ASP -1.SG

‘I will even/also read books’ / ‘will I read books?’

• This element can show up in between an auxiliary and a verb as well, (19)
• Crucially, these particles can also occur ‘inside’ the mismatched cases with a single VH domain
  – *\[ yakal -iyor =mu ol -acak \]
catch -PROG =Q AUX -FUT
catch =Q PROG AUX -FUT
‘Will s/he be catching?’

  – This Y is in fact an auxiliary

(20) a. \[ [ yakal -iyor =mu y -du ]
catch -PROG =Q Y PST
‘Was s/he catching it?’
b. *\[ yakala =mu -yor -di ]
catch -yor -PROG -di -PST

• Thus, particle placement and the domain for stress align.

• Second, It is possible to conjoin two units under a single auxiliary, (21a), but it is not possible to
conjoin two elements under aspectual markers, and have only the bare verbs conjoined, (21b).³

(21) a. \[ [ Yakal -iyor ve yi -yor ]asp ol-acak
catch -PROG and eat -PROG AUX -FUT
‘S/he will be catching and will be eating’
b. *\[ Yakala ve yi ] -yor ol-acak.
catch and eat -PROG AUX -FUT

• It is possible to conjoin the parts that contain stress, but not parts where stress is missing

• What about the mismatched cases? They pattern like the auxiliary cases
  – Conjunction of the units relevant for stress is possible, but nothing smaller:

(22) a. \[ [ gel-'miş ve git-'miş ]-ti-m
come-PRF and go-PRF -PST-1.SG
‘I had come and gone’
(Kornfilt 1996)
b. *\[ gel ve git ]-'miş -ti-m
come and go -PRF -PST-1.SG
‘I had come and gone’

• Thus: coordination, particle placement and stress all align, whereas VH can apply to a bigger unit.

• Coming back to the phonological rules:

(23) **Turkish phonological processes:**
  a. Apply stress to the last syllable of a **syntactic** word
  b. Apply vowel harmony to a **phonological** word

³Some of the coordination data is already discussed in the literature, sometimes in the context of suspended affixation (Kornfilt 1996, Newell 2008, Zanon 2014). In the dissertation I review that data and add missing data to form a complete paradigm.
• I used the same reasoning for Japanese, which yielded the same results
  – Phonological phenomena: Pitch Accent
  – Syntactic phenomena: coordination, particle/adverb placement
• A pitch accent domain coincides with what can be coordinated, or placed ‘inside’ the word

1.3 Feature combinations

• The following list contains TMA markers in Turkish, in the order in which the morphemes occur:

(24) $V_{\text{VOICE -LOW MOD. -ASP}_1 -\text{ASP}_2 -\text{HIGH MOD. -TENSE -MOOD}}$

(25) **Voice**
  a. pass.: *Il*
  b. caus.: *Dur*

(26) **Low mod.**
  a. ability: *abil*

(27) **Aspect 1**
  a. hab.: *Ar*
  b. prog.: *Iyor*
  c. prosp.: *AcAK*

(28) **Aspect 2**
  a. perfect: *miş*

(29) **High mod.**
  a. neces.: *mAll*
  b. possi.: *abil*

(30) **Tense**
  a. future: *AcAK*
  b. anterior: *DI*

(31) **Mood**
  a. evaluative: *sA*
  b. evidential: *miş*
  c. past: *DI*
  d. irrealis *sA*

• Each morpheme can attach to the verb by itself, without any other morphemes (Fenger 2017).
• Voice and Low Modality can be combined with any other morpheme
  – Morphemes in these groups can also be combined with each other
• All morphemes from (27-31) cannot be combined with any other group
  – As soon as aspect is marked on the verb, all other morphemes are auxiliaries
  – There is a cut-off point in (24) after Aspect

(32) `[ bit -ir -i’yor ] [ ol -mali ] [ i -di ]
    finish -CAUS -PROG AUX -NEC AUX -PST
    ‘s/he had to be finishing’

(adapted from Kornfilt 1997)

---

• Crucially, phonology can obscure the syntax. Timing when morphemes are mapped into a ‘word’
  – In the syntax → units are relevant for syntactic operations
  – In the phonology → units are not relevant for syntactic operations

• The exact cut-off point is found in Japanese

• The smaller (syntactic) units that are picked out in two head final languages are similar units picked out in English:

\[
(33) \quad \begin{align*}
  a. & \text{ Ziggy } \left[ \text{play -ed} \right] \text{ guitar (jamming good with Weird and Gilly)} \\
  b. & \text{ *Ziggy } \left[ \text{play -ing -ed} \right] \text{ guitar} \\
  c. & \text{ Ziggy } \left[ \text{was} \right] \left[ \text{play -ing} \right] \text{ guitar}
\end{align*}
\]

1.4 Analysis

What needs to be accounted for is:

• The difference between synthetic and periphrastic forms:

\[
(34) \quad \begin{align*}
  a. & \left[ \text{kal -di} \right] \text{stay -PST} \quad \text{‘s/he stayed’} \\
  b. & \left[ \text{kal -'yor} \right] \left[ \text{i -di} \right] \text{stay -PROG AUX -PST} \quad \text{‘was staying’}
\end{align*}
\]

• The difference between syntactic and phonological words (i.e. stress/VH mismatches)

\[
(35) \quad \left( \left[ \text{kal -'yor} \right] -\text{du} \right) \text{stay -PROG -PST} \quad \text{‘was staying’}
\]

• The cut-off point after *Aspect with regard to syntactic word building

• I assume that there is a constraint against certain heads remaining stranded.
  – One way to resolve a violation is via head movement;
  – If that is not available, auxiliaries can be inserted (Bjorkman 2011, Calabrese 2019)

\[
(36) \quad \text{Affix Constraint: } \left[ \text{INFL} \right]_{x_0}
\quad \begin{align*}
  a. & \text{ Head movement: } \left[ \text{V INFL} \right]_{x_0} \\
  b. & \text{ Auxiliary insertion: } \left[ \text{AUX INFL} \right]_{x_0}
\end{align*}
\]

• One way that head movement can be blocked, in Turkish, is via phases:

\[
(37) \quad \text{Proposal: Phase heads can stop head movement}
\]
• When phase heads stop head movement, any additional word-formation is regulated in the morphology

  – I assume *Aspect* can be part of the inner phase (Aelbrecht 2010, Harwood 2013, 2015, Aelbrecht and Harwood 2015, Wurmbrand 2017)
  – After the phase is sent to Spell-Out, head movement takes place

• Syntax and morphology/phonology are interleaved

  – Unmarked heads are removed (Embick 2010, Calabrese 2019)
  – When the phase head is marked, the word is sent to phonology;
  – When phase head is unmarked, it is removed, and $X^0$ remains in syntax

• Periphrastic verb: PROG-PRS

  – Head movement after spell out (39)
  – Removal of unmarked heads (41), leaves Asp
  – Phase head makes the $X^0$ be interpreted in phonology

(38) Syntax PROG-PST

(39) Head movement, First domain

(40) Removal, First domain

(41) After Removal

– Phonology applies & is fed back into the syntax, summarized in (42).
(42) a. Stress on every last syllable of every highest $X^0$ with a root [Turkish]
b. Apply vowel harmony at every $\omega$

(43) 1. Phase 1 = Asp Turkish
     Morphosyntactic Structure
     
     \begin{align*}
     \text{a. Movement} & \quad [[[[V\ v\ Asp(+\text{PROG})]_{\text{ASP}(+\text{PROG})}]_{\text{ASP}(+\text{PROG})}]_{\text{ASP}(+\text{PROG})}]_{\text{ASP}(+\text{PROG})} \\
     \text{b. Removal} & \quad [[[[V\ ]\ Asp(+\text{PROG})]_{\text{ASP}(+\text{PROG})}]_{\text{ASP}(+\text{PROG})}]_{\text{ASP}(+\text{PROG})}
     \end{align*}
     
     Phonological grouping
     
     c. vocabulary insertion $[[/\text{kal}/\ V\ /\text{iyor}/\ ]_{\text{ASP}(+\text{PROG})}]_{\text{ASP}(+\text{PROG})}$
d. stress assignment $[[/\text{kal\ iyor}/\ ]_{\text{ASP}(+\text{PROG})}]_{\text{ASP}(+\text{PROG})}$
e. vowel harmony $[[/\text{kal\ iyor}/\ ]_{\text{ASP}(+\text{PROG})}]_{\text{ASP}(+\text{PROG})}$

– second cycle: T is stranded $\rightarrow$ AUX is inserted
– Movement is not possible anymore

(44) Second domain, Stranded heads
     (kali‘yor)$_\omega$ (idi)$_\omega$

- Synthetic verb (simple tense)
  – The derivation is similar to periphrastic tense, movement applies
  – However, during removal (46), the phase head is removed because unmarked
  – *The syntactic $X^0$ is not closed and is fed into the syntax*, before phonological interpretation (47).
    o Tense can then attach to the verb.

(45) Movement, First domain
     (46) After Removal
     (47) Movement Past Tense

\begin{align*}
\text{AspP} & \quad \text{vP} & \quad \text{Asp} & \quad \text{VP} & \quad \text{t}_v \\
\text{Asp} & \quad \text{Asp} & \quad \text{VP} & \quad \text{t}_v & \quad \text{v} & \quad \text{V} & \quad \text{[PST]} \\
\text{V} & \quad \text{t}_v \\
\text{[PST]} & \quad \text{T} \quad \text{V} & \quad \text{T}
\end{align*}
Phases can delimit word-formation processes. This predicts …

1. Timing differences have an effect in other locations (Sec. 2, 3)
   (a) HM takes place after Spell-Out, and correlate with other phenomena (Sec. 2)
   (b) T and V are separate in the syntax (Sec. 3)

2. Phases interact with word building (Sec. 4)
   (a) The form of V should correlate with other phasal phenomena (Sec. 4)
   (b) Morphemes below the phase head should not participate in periphrasis (Sec. 4)

1.5 Words within words

• What needs to be accounted for is, why stress and vowel harmony sometimes do not align\(^5\)

• What is key is when material is put together and when phonological operations apply.
  – I propose this can be derived via Movement of PST after phonological material is spelled out.

• The derivation that Shwayder (2015) proposes is abstractly represented in (48).\(^6\)

(48) Move with Interleaved Phonology (Shwayder 2015, p.211)

<table>
<thead>
<tr>
<th>Morphosyntactic Structure</th>
<th>[ (/X/) _X _Y ]_0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Move</td>
<td>[ (/X/) _X _Y ]_0</td>
</tr>
<tr>
<td>Phonological Grouping</td>
<td>[ (/X/) _X _Y ]_0</td>
</tr>
</tbody>
</table>

• Important: takes place after the first X\(^0\) is spelled out, after stress assignment.

• The derivation for Turkish is as follows

(49) \((( \text{kal } -i\text{yor }) -\text{du })\)
    stay -PROG -PST
    ‘was staying’

  – Morphosyntactic structure: Two X\(^0\), main verb and Tense
  – Movement takes place after this operation
  – Stress assignment does not take place because there is no lexical root anymore
  – The auxiliary deletes, since it occurs between two consonants,
  – Vowel Harmony takes place → spreads from verb

\(^5\)The analysis for Turkish is in line with the one proposed by Kornfilt (1996).
\(^6\)For ease of exposition, I have omitted some of the notations which indicates if something is linearized or not. What is important for the discussion at hand is how a syntactic word can form a recursive phonological structure.
Morphosyntactic structure

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Auxiliary insertion</td>
<td>[ (/kal’iyor/ )_ω ASP(+PROG) [AUX T ]_T ]</td>
</tr>
<tr>
<td>b. Move</td>
<td>[ (/kal’iyor/ )_ω AUX T ]_T</td>
</tr>
</tbody>
</table>

Phonological grouping

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Auxiliary Insertion</td>
<td>[ (/kal’iyor/ )_ω /y/ /DI/ ]_T</td>
</tr>
<tr>
<td>b. Stress</td>
<td>n/a</td>
</tr>
<tr>
<td>c. Auxiliary form</td>
<td>[ ( /kal’iyor/ )_ω /DI/ )_ω ]</td>
</tr>
<tr>
<td>d. Vowel Harmony</td>
<td>( ( /kal’iyor/ )_ω /du/ )_ω</td>
</tr>
</tbody>
</table>

- Units where syntax and phonology align → syntactic words built via head movement
- Mismatches (aka ‘words within words’) are created after vocabulary insertion
  - Crucially, this operation does not have an effect on the syntax anymore

2 Arriving (very) early: Matching differences in VSE

- Part of the analysis is that Move>SO. However, Move<SO has also been proposed
- Luckily, there are tests that distinguish between before/after spell-out: Ellipsis
- I show that Japanese and Turkish patterns with languages that have head movement after SO
- Languages differ in what can be elided and what can be pronounced
  - English: can elide a VP and pronounce do, (51);
  - Irish: can elide a VP and pronounce V, (52)

(51) Posh Spice wanted to **buy** a black Gucci dress, and she **did** buy a black Gucci dress

(52) a. Ar **chuir** tú isteach ar an phost?
    ‘Did you apply for the job?’

b. **Chuir** / Níor **chuir**
    ‘I did / I didn’t’ (McCloskey 2011)

- Stranding a V, and eliding a VP is called **Verb Stranding Verb Phrase Ellipsis** (VSE)
- What is of interest here are the requirements on the stranded V.  
  - … require the antecedent and stranded verbs to be **identical**, (53)
    - Irish, Scottish Gaelic, Uzbek
  - … can have a **different** verb for the antecedent and stranded V, (54)
    - Russian, Hungarian, European Portuguese, Swahili, Greek, Hebrew

---

7Discussed in the literature as part of the **Verbal Identity Condition** (Goldberg 2005).
(53) *Níor **cheannaigh** mé teach ariamh, ach dhíol.
NEG.PST buy I house ever but sold
‘I never bought a house, but I sold one’ (McCloskey 2017, 128)

(54) a. Našël li Paša knigu v biblioteke, i žurnal v stolovoj?
find.PST ‘SG.M Q Paša book.ACC in library.PREP and magazine.ACC in cafeteria.PREP
‘Did Pasha find a book in the library, and a magazine in the cafeteria?’

b. Net, ne Našël, a poterjal
No, NEG find.PST ‘SG.M but lose.PST ‘SG.M
‘No, he didn’t find ( . . . ), he lost ( . . . )’

• This difference has been argued to come from the timing of head movement (Gribanova 2018, 2020)
  – When HM occurs *before* Ellipsis matching (E), the verb is not part of the VP
  – When HM occurs *after* E, the verb is ‘trapped’ when matching takes place

(55) a. Step #1: Movement

```
FP
  
  F  VP
    
      F  V  ADV  VP
         V  O
```

b. Step #2: Identity

```
FP
  
  F  VP[E]
    
      ADV  VP
         V  O
```

(56) a. Step #1: Identity

```
FP
  
  F  VP[E]
    
      ADV  VP
         V  O
```

b. Step #2: Movement

```
FP
  
  F  VP[E]
    
      ADV  VP
         V  O
```

• Even though both Russian and Irish have VSE, they differ on identity requirements of the verbs
  – In my system, this means HM is timed differently: before or after Spell-Out

• Prediction (for Head Final Languages):
  – If HM happens before SO ⇒ no Verbal Identity
  – If HM happens after SO ⇒ Verbal Identity
2.1 Head movement in Head final languages

- Does Japanese have VSE? It does have something elided . . .

(57) Totoro-wa booru-o ut-ta kedo, Mei-wa booru-o ut-anak-atta
Totoro-TOP ball-ACC hit-PST but Mei-TOP ball-ACC hit NEG-PST
‘T. hit ball, but Mei didn’t hit ball’ based on (Funakoshi 2016, 114)

- Three analyses have been put forward:

(58) Possible analyses for (57) (from Funakoshi 2016, 114)
   a. Mei-TOP pro hit-NEG-PST pro-analysis
   b. Mei-TOP baru-o hit-NEG-PST Argument ellipsis-analysis
   c. Mei-TOP [VP baru-o t] hit-NEG-PST VSE-analysis

- It is well established that Japanese does have pro.

- There are many constructions where argument ellipsis seems a better analysis than VSE (as observed by Funakoshi 2014, 2016, Sakamoto 2017).
   → This has often been taken as evidence that Japanese does not allow VSE at all.

- Solution? Adjuncts

- The adjunct reading is available with the correct context (59) (Takahashi 2008, Funakoshi 2014, 2016).
  - Add context
  - Second clause negated
  - use ‘but’ or ‘also’.

(59) Context: Pikachu and Togepi washed their owner’s cars to get food. Pikachu was thorough in his work, while Togepi was not. (based on Funakoshi 2016:119)
   a. Pikachu-wa kuruma-o teineini arat-ta
      Pikachu-TOP car-ACC carefully wash-PST
      ‘Pikachu washed the car carefully’
   b. Togepi-wa kuruma-o teineini arat-anak-atta.
      Togepi-TOP car-ACC carefully wash NEG-PST.
      Togepi-ga arat-ta atto-no kuruma-wa kitanak-atta
togepi-NOM was-PST after-GEN car-TOP dirty-PST
      ‘Togepi did not wash the car carefully. The car that Togepi washed was dirty’

8The same problem arises in Turkish, but taking into account the various tests, Turkish shows the same result as Japanese.
• **Conclusion:** Japanese does have VSE … So what about identity?

• We can use the above test. And then… verbs cannot mismatch (based on Funakoshi 2014, 331-2).

(60) *Togepi-wa (itumo) zikandoorini ku-ru kedo,
    T.TOP always on.time come-PRS but
Ookido-hakase-wa (itumo) zikandoorini omicninar-ana-i
Oak.-prof.-TOP always on.time come.HON-NEG-PRS

Intended: ‘Togepi always comes on time, but Prof. Oak always doesn’t come on time’

• This means there are two types of languages.
  – Head movement before Spell-Out (Greek, Russian)
  – Head movement after Spell-Out (Irish, Japanese, Turkish)

---

Phases can delimit word-formation processes. This predicts …

1. Timing differences have an effect in other locations (Sec. 2, 3)
   (a) √ HM takes place after Spell-Out, and correlate with other phenomena (Sec. 2)
   (b) T and V are separate in the syntax (Sec. 3)

2. Phases interact with word building (Sec. 4)
   (a) The form of V should correlate with other phasal phenomena (Sec. 4)
   (b) Morphemes below the phase head should not participate in periphrasis (Sec. 4)

---

3 **Towards a typology of word-building**

• Word building has been discussed a lot in the literature:
    ○ *phase extension* (den Dikken 2007, Gallego and Uriagereka 2007b,a)
    ○ *At PF*: (Chomsky 2001, Schoorlemmer and Temmerman 2012, Platzack 2013)
  – **Phonology:** *Local Dislocation*  (Embick and Noyer 2001, Embick 2007, Shwayder 2015)

• What if it is possible to parametrize when movement takes place.² Merge two heads in the
  – **Syntax:** Before or after (phased) spell-out: Phono=Morpho=Syn
    ○ domain extension, different constraints on ellipsis

²This approach is building on approaches in the literature that argue that there are various locations as well
- **Morphology**: Phono = Morpho ≠ Syn
  - Synthetic forms might behave like periphrastic forms in the syntax
- **Phonology**: Phono ≠ Morpho ≠ Syn
  - Periphrastic forms and mismatched forms should be syntactically the same
  - Synthetic forms should behave differently

### Start

### Syntax

### Morphology

### Phonology

### Output

**Table 2: Movement at different stages**

- The idea is to not just analyze ‘words’ via a specific mechanism, but see if this mechanism then matches up with what we should predict in different modules
- Looking at word building and mismatches in this way is also useful for when the output masks the underlying structure, i.e. O1 vs O3, and verb complexes in head final languages.

#### 3.1 Arriving fashionably late: morphology vs syntax

- Can it be the case that the output is the same, but the derivation is different?
Table 3: Movement at different stages

- Many have looked at correlations between the height of the verb and the morphological form.

(61) Scary Spice \( \times \{ \text{kisses} \} \) often \( \checkmark \{ \text{kisses} \} \) Baby Spice \[ \text{[English]} \]
(62) Obelix \( \checkmark \{ \text{embrasse} \} \) souvent \( \times \{ \text{embrasse} \} \) Idéfix
Obelix kisses often kisses Dogmatix
‘Obelix often kisses Dogmatix’ \( \text{[Pollock 1989, adapted p.367]} \)

- What is of interest to me is whether the verb in both of these languages is built in the same way
  - In V-to-T movement languages, it has been argued that the verb is formed through head movement
  - In a language like English this is done through a morphological operation (affix hopping, morphological merger, lowering..)
  - Or differently: Lasnik (1995) argues that English verbs are built in the syntax, whereas French verbs are built in the lexicon

(63) V-to-T movement

\[
\begin{align*}
\text{TP} & \quad \text{T} \quad \text{VP} \\
\text{T} & \quad \text{V} \quad \text{T} \quad \text{ADV} \quad \text{VP} \\
\text{V} & \quad t \quad \text{XP}
\end{align*}
\]
(64) No V-to-T movement

a. Syntax

```
TP
  T     VP
    ADV   VP
      V    XP
```

b. Morphology

```
TP
  t    VP
    ADV   VP
      V    XP

V    T
```

- Put differently: In order to make the English simple tense, an additional step is needed
  - That is, in the syntax T and V remain separate, whereas in V-to-T movement languages it is not
- Thus there is theoretically speaking a mismatch between syntax and morphology:

(65) a. Syntax: [V] [T]

b. Morphology: [V -T ]

- So, in the syntax, English simple tenses should behave separately, which is similar to periphrastic constructions:
  - Even though there is a way for V+T to end up in a single morphological word, there is a step in the derivation where both the progressive T and simple T are the same.

(66) TP

```
TP
  T     VP[E]
    V    XP
```

- This can be tested by looking at inflectional mismatches in VP-ellipsis (Sag 1976, Warner 1986, Lasnik 1995)
  - If VP ellipsis happens at the point of spell-out, elements outside of the cycle should not be part of the matching condition
  - Specifically, if it happens when T and V are not together, T should be able to mismatch, (67)
  - The progressive on the other hand is not allowed to mismatch

(67) Scary Spice slept, and Posh Spice will too

a. *Scary Spice slept, and Posh Spice will slept too
b. Scary Spice slept, and Posh Spice will sleep too

adapted from Lasnik (1995)

(68)

\[
\begin{array}{c}
\text{TP} \\
\downarrow \\
\text{Scary Spice} \\
\downarrow \\
\text{T} \\
\downarrow \\
\text{VP} \\
\downarrow \\
[\text{PST}] \\
\downarrow \\
\text{V} \\
\downarrow \\
\text{sleep}
\end{array}
\]

• On the other hand, the prediction is that for languages that have movement earlier on, there should be no such mismatches.
  – In this case the V+T complex is already built in the syntax, before spell-out and before ellipsis matching has taken place.

• But where to look? Many languages do not have this type of ellipsis.
  – Maybe European Portuguese?
  – This language allows for VP ellipsis when the constituents match, (69)
  – When the finiteness of V differs, mismatches are not allowed, (70)

(69) O João já tinha lido este livro, mas a Maria não tinha lido este livro
the João already had read.PRT this book but the Maria not had read.PRT this book
‘João had already read this book, but Maria hadn’t (read this book)
(Nunes and Zocca 2009, 33)

(70) a. *Ela perguntou se alguém lera o jornal,
she asked if anybody read.PLF the newspaper
mas ninguém tinha lido o jornal
but nobody had read.PRT the newspaper
‘She asked if anybody read the newspaper, but nobody had (read the newspaper)
(Cyrino and Matos 2005, 14)

b. *A Maria estudou muito, mas o João não vai estudar muito
the Maria studied much but the Joao not goes study.INF much
‘Maria studied very hard, but Joao will not (study very hard) (Todorović 2016, 28)
Phases can delimit word-formation processes. This predicts …

1. Timing differences have an effect in other locations (Sec. 2, 3)
   (a) ✓ HM takes place after Spell-Out, and correlate with other phenomena (Sec. 2)
   (b) ✓ T and V are separate in the syntax (Sec. 3)

2. Phases interact with word building (Sec. 4)
   (a) The form of V should correlate with other phasal phenomena (Sec. 4)
   (b) Morphemes below the phase head should not participate in periphrasis (Sec. 4)

4 Restrictions on word building

- One part of the analysis is the relation between word building and phases, as a way of limiting when words can be built (in the syntax)

- Theoretically speaking two options have been proposed:
  - Phases inside words
  - Phases delimit words

(71) a. \[z\]
     b. \[z\]

(72) If Darth Vader says that Han Solo was being frozen in carbonite, then ...
    a. [being frozen in carbonite] he was.
    b. *[frozen in carbonite] he was being. (Harwood 2015:550, 63)

(73) a. *Someone told Ziggy to play the guitar, and [play-ed the guitar] he
b. Someone told Ziggy to play the guitar, and [play the guitar] he did

- The same holds in Japanese for Fronting:
  - Aspect morpheme can be fronted; Tense morpheme cannot
  - Verb stem can be fronted, which is similar to English (both with do support)

\[\text{(74) a. } \left[\text{ASPP aogaeru-} \text{tabe-te-sae/mo/wa/dake } \right] \text{ Kaonashi-ga } t_{\text{ASPP}} \text{ i-ta } \left[\text{Asp} \right] \\
\text{Aogaeru-ACC eat-ASP-even/also/TOP/only } \text{No.Face-NOM } \text{AUX-PST} \\
\text{‘No Face was even/also/only eating Aogaeru’}\]

\[\text{b. } \left[\text{TP aogaeru-} \text{tabe-ta-sae/mo/wa/dake } \right] \text{ Kaonashi-ga } t_{\text{TP}} \left[\text{Tense} \right] \\
\text{Aogaeru-ACC eat-PST-even/also/TOP/only } \text{No.Face-NOM} \]

\text{‘No Face even/also/only ate Aogaeru’}\]

- Other phenomena looked at: particle placement in Japanese (Topic) and Turkish (Q-particle)
- Ellipsis & interpretation of Adverbs (Japanese)
  - This also bears on the question of what is sent to spell-out: whole phase or the complement

- **Prediction # 2**: Variation (33 lgs, 18 fam) with regard to expression of Tense, Mood, Aspect:
  - Head movement is blocked by a phase;
  - Morphemes ‘inside’ the phase can move, no need for AUX

\[\text{(75) TP} \]

\[\text{T PerfP}\]

\[\text{Perf AspP}\]

\[\text{Asp VoiceP}\]

\[\text{Voice Lex.AsP}\]

\[\text{Lex.AsP vP}\]

\[\text{v VP}\]

\[\text{V XP}\]

- The focus was on looking at which TMA combinations are (im)possible in a single X\(^0\). Two variables:
  - What are the inflectional categories in the languages?
- What is the status of the morpheme: affixal or free morpheme?
  
  o Turkish T: affixal, and needs a host:

  (76) a. \[ kal \ -\ '{\text{di}} \ ]
      stay -PST
      \text{‘s/he stayed’}
  
  b. \[ kal \ -{\text{iyor}} \ ]\ [ i \ -\ {\text{di}} \ ]
      stay -PROG AUX -PST
      \text{‘was staying’}

  o Japanese and English: free morphemes for \textit{begin}, (78, 77), but differ in one or two words.

  (77) Lyra \textit{begins to become distant} from Pantalaimon on the way to the Land of the Dead

  (78) \[ hedatar \ -i \ -{\text{hajime}} \ -ru \ ]
      get.distant -INF -begin -PRS
      \text{‘begin to become distant’}

- \( v \), valency changing and lexical aspect categories can only be periphrastic, when expressed as a root (i.e. \textit{begin, make}), but not as a functional morpheme

\textbf{(79) Generalization I}

When \( v \), Valency changing morphology, and Lexical aspect are expressed as a functional morpheme, they do not participate in periphrasis

- In principle it could be expected that any morpheme can be a periphrastic construction,

  - so why is it that certain categories when they are affixal never do?
  - I.e., why don’t we find alternations as in (77), just as for past tense in Turkish?

\textbf{(80) Not attested:}

a. \( [V \ -\text{CAUS}] \) vs. \( [V-x] \ [AUX \ -\text{CAUS}] \)

b. \( [V \ -\text{INGRESSIVE}] \) vs. \( [V-x] \ [AUX \ -\text{INGRESSIVE}] \)
5 Conclusion and outlook

- I argued that it is possible to maintain that there is a coherent unit in grammar we call a word.

- In order to show this, I have looked past the messiness:
  1. Mismatches between phonology and syntax
     - Units where syntax and phonology align are syntactic words
     - Mismatches come from timing differences in putting the relevant elements together
  2. Timing of Head Movement
     - Consequences for verbal identity in ellipsis
     - Typology of timing
  3. Limits on word building defined in terms of phase hood
     - Correlations between words and phrases
     - Consequences for which features participate in periphrasis

- Methodological Contribution:
  - Cross-linguistic comparison
  - Cross-module comparison

References

Bošković, Željko. (2014). Now I’m a phase, now I’m not a phase. LI 45(1), 27–89.
Mouton de Gruyter.