Coarticulation as an *epiphenomenon* of syllable-synchronized target approximation

—Evidence from $F_0$-aligned formant trajectories in Mandarin

Hong GAO
Sichuan University, China

Yi XU
University College London, UK
Classical V-V anticipatory coarticulation

- Öhman (1966): V-to-C F2 transition varies in direction depending on the vowel of the following syllable

- (Öhman 1966:165): “a motion toward the final vowel starts not much later than, or perhaps even simultaneously with, the onset of the stop-consonant gesture”

- But no conclusion on this point later in the paper. And a vowel gesture is believed to start sometime later than the consonant gesture in CV.

- No subsequent systematic comparison of C-V alignment to our knowledge.
How can we know the onset and offset of a segment?

Strategy 1: Use $F_0$ alignment as independent time reference


- **Note**: The reverse should also be true: Segments events are consistently aligned with $F_0$ turning points (Xu & Liu, 2007; Gao & Xu, 2010).
• **Strategy 2:** Compare continuous formants in minimal pairs, treating movement departure and convergence points as articulatory onset and offset

![Graph showing F0 peaks indicating conventional syllable boundaries](image)

- **F0 peaks indicating conventional syllable boundaries**
- **F0 (Hz)**
- **(F2+F3)/2**
- **Normalized time**

- ni yi wei
- ma yi wei
- mao yi wei
- lou yi wei
Why \((F2+F3)/2\)?

- Affiliation of F2 and F3 often shows quantal shifts (Stevens 1972; Stevens & Keyser 2010)

Mean formant tracks of 10 repetitions by a male speaker
Stimuli & procedure

1. Formant trajectory from $V_1$ to $V_2$, across intervening /l/:
   Minimal contrast in V-to-V movement: $ni$-$lu$ / $lu$-$li$ (倪庐/卢黎)

2. Relative onset of movements toward $C_2$ and $V_2$:
   Minimal contrast in C-onset: $ni$-$li$ / $ni$-$ji$ (倪黎/倪姨)
   Minimal contrast in V-onset: $ni$-$li$ / $ni$-$lu$ (倪黎/倪庐)

✧ All words have R-R tone sequence to give 2 $F_0$ peaks

✧ 3 male and 4 female speakers of Beijing Mandarin, reading the material at normal speed, with 8 repetitions
Results 1: $F2$-3 movements toward V2 extended continuously from the center of V1 to the center of V2, across the intervocalic [i]

✧ Movement toward V2 both *starts* and *ends* earlier than conventional spectral landmarks
Target Approximation model
(originally proposed for tone) (Xu & Wang, 2001)

- These formant movements resemble pitch movements toward a static underlying tonal target
Results 2: F2-3 movements toward C2 & V2 start simultaneously, from center of V1

C onsets per l/j contrast

V onsets per i/u contrast

ni li / ni ji

ni li / ni lu

lei li / lei ji

lei li / lei lu

ma li / ma ji

ma li / ma lu
**Statistical comparison of onset of divergence in C and V minimal pairs**

1. Running t-test: Onset of divergence in each minimal pair = the time at which $p$ drops and remains below 0.05
2. Repeated measures ANOVA: Whether C and V minimal pairs differed in onset of divergence
3. Results: No significant difference between C- and V-divergence but significant difference in divergence onset between vowels.
Discussion

- Assuming that articulating a segment is to *approach its underlying target*,
- Assuming also that coarticulation is *concurrent articulation of multiple segments*,
- Then, there is no anticipatory co-articulation of V2 with V1, because the articulation of V1 is already finished when the articulation of V2 starts
- There is also no anticipatory coarticulation of V1 with C2, because, again, the articulation of V1 terminates as the articulation of C2 starts
- But there is genuine coarticulation between C2 and V2
The syllable specifies the temporal alignment of all the constituent phones, including C, V, T and P, under 3 principles:

- **Co-onset** — Initial C, first V, T and P all start at the syllable onset.
- **Sequential offset** — Non-initial segments, whether V or C, are sequentially aligned after the first V of the syllable.
- **Synchrony of laryngeal phones** — Both T and P are synchronized with the entire syllable to which they are associated.
Conclusion

- Articulation of C and V in a syllable start at about the same time, and well before the landmark-based syllable boundary.

- Articulation of V terminates well before the landmark-based syllable boundary.

- Overall, genuine CV co-production occurs only between onset C and the following V, while the rest of the “coarticulation” is only an epiphenomenon of syllable-synchronized target approximation.
An experiment was carried out to test the hypothesis that the syllable is a time structure that synchronizes tonal, consonantal and vocalic target approximation movements.

The strategy was to align formant movements with $F_0$ turning points of lexical tones as time reference, and then assess the temporal scope of articulatory movements by comparing formant trajectories and their turning points across minimal pairs.

Native Mandarin speakers produced C1V1#C2V2 disyllabic sequences where C2 is /y/, /w/ or /l/, and V1 and V2 varied in height and frontness. Analysis of $F_0$-aligned F2-3 (average of F2 and F3) trajectories revealed patterns in support of the main hypothesis.

First, movements clearly discernable as approaching either C2 or V2 targets started at about the same time from the center of V1, i.e., well before the conventional landmark-based syllable boundary.

Second, some F2-3 trajectories extended continuously from the center of V1 to the center of V2, across the intervening /l/, indicating a long and uninterrupted V2 approximation movement.

These results provide support for the view that genuine CV co-production occurs only between onset C and the following V, while the rest of the “coarticulation” is only an epiphenomenon (arising from landmark-based segmentation) of syllable-synchronized target approximation.