

# On the dynamic behavior of Vowel-to-Vowel Harmony in French: Do speakers control states or changes?

Olivier Crouzet & Agnieszka Duniec

Laboratoire de Linguistique de Nantes (LLING – EA3827)  
Département de Sciences du Langage  
UFR Lettres et Langages  
Université de Nantes

October 29, 2014– 1st DINAFON Meeting – UNICAMP



# Plan de la présentation

- 1 Vowel harmonisation in French: Background
  - Classical accounts
  - Acoustic data and temporal behavior
- 2 Further Issues
  - Coarticulatory transitions?
  - Monolithic vs. dynamic control targets
- 3 Investigation of the impact of vowel duration on V-to-V Harmony
  - Aims
  - Procedure



# Plan de la présentation

- 1 Vowel harmonisation in French: Background
  - Classical accounts
  - Acoustic data and temporal behavior
- 2 Further Issues
  - Coarticulatory transitions?
  - Monolithic vs. dynamic control targets
- 3 Investigation of the impact of vowel duration on V-to-V Harmony
  - Aims
  - Procedure



# Plan de la présentation

- 1 Vowel harmonisation in French: Background
  - Classical accounts
  - Acoustic data and temporal behavior
- 2 Further Issues
  - Coarticulatory transitions?
  - Monolithic vs. dynamic control targets
- 3 Investigation of the impact of vowel duration on V-to-V Harmony
  - Aims
  - Procedure



## Vowel-to-Vowel Harmonisation / assimilation in French

- It has first been described by Grammont (1933);
- A regressive assimilation of the properties of a vowel over the preceding vowel, through an intermediate consonant, within a word;
- It would affect both aperture and anteriority;

/eme/  
[eme]

(fr. “to love”)

/emabl/  
[ɛmabl]

(fr. “kind,  
friendly”, same  
radical as “to  
love”);



## Behavior

- Same “underlying vowel” ( $V_1$ );
- Gives rise to an alternation between vowels with various degrees of aperture;
- Would only apply to mid vowels  $\{e/\epsilon, \emptyset/\text{œ}, o/\text{ɔ}\}$ ;
- Depending on the following vowel context ( $V_2$ );
- Speakers would alternate between the following forms. . .

[eme]

[εma]



## Behavior

- May give rise to only slight spectral variations;
- Usually not noticed by the speakers or listeners;
- Preventing from producing it *does not seem to be ungrammatical* at all and is unnoticed by speakers;
- May be interpreted in terms of a long-term coarticulation effect rather than in terms of vowel harmony *per se*;
- Therefore seems to differ fundamentally from (e.g.) Turkish or Hungarian (Magyar) VH;



## Subjective vs. Objective analyses

- Most discussions have been based on subjective auditory judgements (Carton, 1974; Dell, 1972; Fouché, 1956; Grammont, 1933; Martinet, 1945, among others);
- Only recently have authors started investigating this phenomenon on the basis of *objective* speakers' productions;
- Main work : Nguyen, N., & Fagyal, Z. (2008). Acoustic aspects of vowel harmony in French. *Journal of Phonetics*, 36(1), 1–27;





## Nguyen & Fagyal (2008)'s main results

- Measured  $V_1$  formant frequencies at the vowel mid-point;
- Confirm classical subjective judgments;
- Both F1 and F2 formant frequencies of  $V_1$  are influenced by the properties of  $V_2$ ;



## Further issues

- Isn't this phenomenon simply an artifact emerging from formant transitions within the V-to-V sequence?
- Is this phenomenon active over the whole  $V_1$ ? Could this help us decipher what the speaker's target is and how it is controlled?



## Why do these should matter?

- It may simply emerge from the articulatory / acoustic trajectory between 2 *unmodified targets* = artifact;
- It may be associated to either (1) a monolithic view of this V-to-V harmony ( $V_2$  “as a whole single target” influences the preparation of  $V_1$  “as a whole single target”) or (2) a dynamical view according to which this phenomenon may evolve over time within  $V_1$ ...



## V-to-V assimilation vs. Coarticulatory artefact

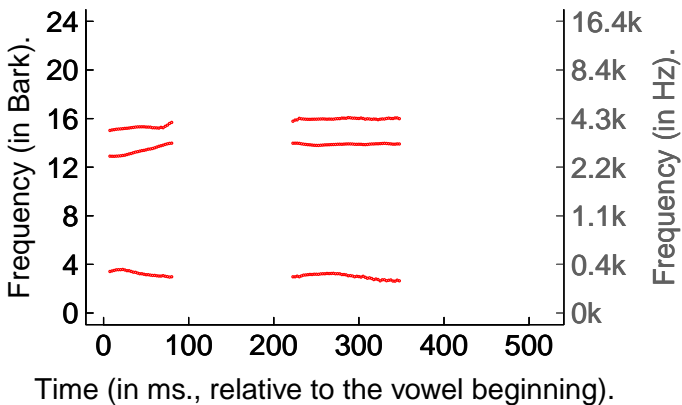


Figure: Example of formant data extracted from a  $V_1CV_2$  sequence, [eki] from the word “béquille”, [bekij], en. “crutch”.

## Note

- Only schematic formant tracks are presented graphically;
- These were extracted with Praat formant tracking and then processed within R;
- They constitute our base data for analyses;



## Formant movements in Vowel Harmony

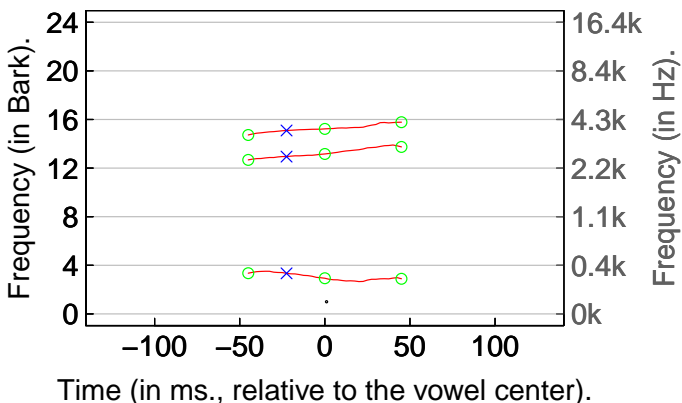


Figure: Zooming on  $V_1$  inside the very same sequence. Formant movement: [e] towards [i]. The blue cross was our alternative hypothetical target (based on maximal  $F_1$  frequency).

## Why do these should matter?

- It may simply emerge from the articulatory / acoustic trajectory between 2 *unmodified targets* = artifact;
- It may be associated to either (1) a monolithic view of this V-to-V harmony ( $V_2$  “as a whole single target” influences the preparation of  $V_1$  “as a whole single target”) or (2) a dynamical view according to which this phenomenon may evolve over time within  $V_1$ ...
- ... which may have implications for modelling articulatory control in this situation;



Does it hold over the whole  $V_1$ ?

Rather strong movement within the vowel

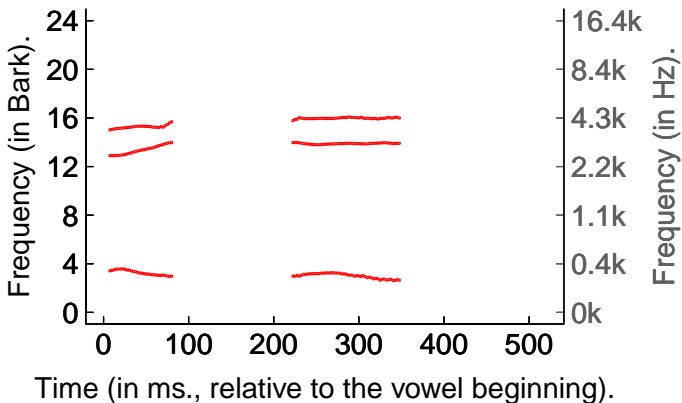


Figure: Example of formant data extracted from a  $V_1CV_2$  sequence, [eki] from the word “béquille”, [bekij], en. “crutch”



Does it hold over the whole  $V_1$ ?  
Almost no movement within the vowel

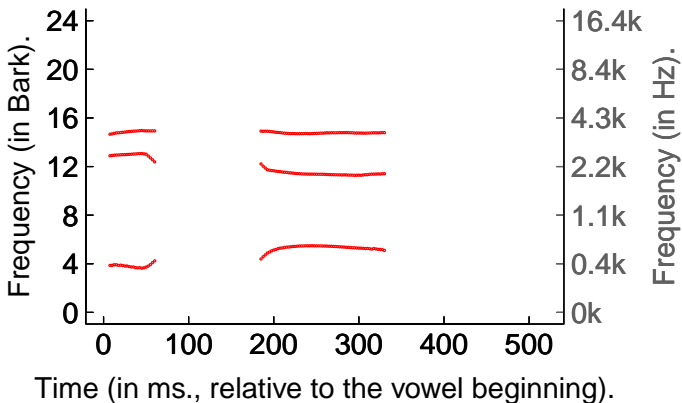


Figure: Example of formant data extracted from a  $V_1CV_2$  sequence, [eta] from the word “bétasse”, [betas], slang for eng. “dumb girl”

## Recent data

- Isn't this phenomenon simply an artifact emerging from formant transitions within the V-to-V sequence? (Duniec & Crouzet, 2013, at PaPI);
- It does not seem so, as the influence of  $V_2$  on  $V_1$  manifests itself even if we change our criterion accordingly for locating the  $V_1$  target;
- Similar significant effects at both positions: vowel mid-point and maximal F1 point;
- Amplitude of the effect 0.20/0.30 Barks on both F1, F2 & F3;
- Is this phenomenon active over the whole  $V_1$ ? Could this help us decipher what the speaker's target is and how it is controlled? — It does appear at the very beginning of  $V_1$  (Duniec & Crouzet, 2014, at LabPhon) but we may need more subtle tools and methods to address this issue. . .



## Recent data

- Is this phenomenon active over the whole  $V_1$ ? Could this help us decipher what the speaker's target is and how it is controlled?
- It does appear at the very beginning of  $V_1$  (Duniec & Crouzet, 2014, at LabPhon)
- Similar significant effects at both vowel mid-point, maximal F1 point and vowel onset point;
- But we may need more subtle tools and methods to address this issue...



## Relationship to $V_1$ duration

- Though we found that significant effects occurred both at mid-point and start-point;
- We may have missed specific cases by not contrasting analyses based on  $V_1$  duration;
- For short  $V_1$ s, the speaker may need to apply VH early;
- Whereas for long  $V_1$ s, he/she may trigger the process later;
- But if speakers control VH as a monolithic “object”, durations should not modify our previous observation;
- VH should then occur as soon as at the vowel start-point no matter the duration;



## Relationship to $V_1$ duration

- Though we found that significant effects occurred both at mid-point and start-point;
- We may have missed specific cases by not contrasting analyses based on  $V_1$  duration;
- For short  $V_1$ s, the speaker may need to apply VH early;
- Whereas for long  $V_1$ s, he/she may trigger the process later;
- But if speakers control VH as a monolithic “object”, durations should not modify our previous observation;
- VH should then occur as soon as at the vowel start-point no matter the duration;



## Data collection

- Septentrional speakers from the Nantes (Britanny) area;
- Same raw data as those investigated in Duniec & Crouzet (2013, at PaPI) and Duniec & Crouzet (2014, at Labphon);
- Alternate analysis based on  $V_1$  vowel duration;
- As for now, 3 speakers have been transcribed, who have produced 160 sequences twice (overall 320 sequences per speaker);



## Word selection process

- bi-syllabic words;
- may or may not share a common radical (no difference according to data analysed by Nguyen & Fagyal, 2008) – approx. 30 out of 80 pairs share a common radical;
- Paired with respect to intermediate consonant (within the  $V_1CV_2$  internal sequence);
- Paired with respect to the (facultative) pre- $V_1$  consonant or consonant cluster;



## Sample material

Example word pair	V1	V2	Sample size
terraine /tɛʁin/ - terrasse /tɛʁas/	Front	i/a	30
ailé /Ele/ - ailette /Elɛt/	Front	e/ɛ	24
dévot /dɛvo/ - dévote /dɛvɔt/	Front	o/ɔ	6
prêteuse /pʁɛtøz/ - prêteur /pʁɛtœʁ/	Front	ø/œ	12
rosine /ʁOzin/ - rosace /ʁOzas/	Back	i/a	28
noter /nOte/ - notaire /nOteʁ/	Back	e/ɛ	32
auto /Oto/ - automne /Otɔn/	Back	o/ɔ	8
donneuse /dOnøz/ - donneur /dOnœʁ/	Back	ø/œ	18

Table: Example word pairs, V1/V2 categories and corresponding sample size; V1 is transcribed as neutral archiphoneme E (Front V1) or O (Back V1)





# Procedure

- Each word is pronounced twice in a constant carrier sentence;
- The whole sentence is displayed on a screen;
- [ildi WORD dɔ̃fwa WORD]
- e.g. [ilditɛvɛ̃ndɔ̃fwa|tɛvɛ̃in]
- In the middle and at the end of the sentence;
- Only the middle occurrence is analysed;
- Although it is not mentioned to the speakers, it is assumed this should tend to prevent them from inserting a pause around the (first) target word;
- The sentences are presented at a comfortable but time constrained pace (one sentence every 1500 ms);



# $V_1$ duration computation / classification

- Computation based on beginning and end of Praat markers for transcription;
- Observations were categorized into 3 duration groups: “short”, “medium”, “long” based on similarly sized quantiles (1/3rd each);
- Only the two extreme groups have been entered into the analysis;
- The short vowels are shorter than 50ms; while the long vowels are longer than 65ms;



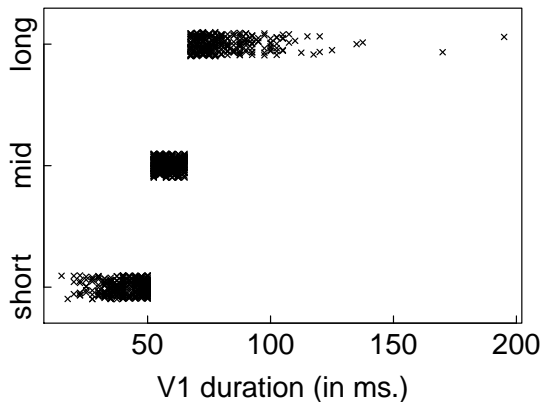
Distribution of  $V_1$  durations depending on quantiles.

Figure: Distribution of  $V_1$  durations depending on quantiles. Each quantile contains 33% of the sample.



Results: t-tests (both front & back  $V_1$ s, to be investigated further

Only 1st formant comparisons do reach significance.

	Short $V_1$ s ( $\leq 50ms$ )	Long $V_1$ s ( $> 65ms$ )
Mid-point	<b>p&lt;.01</b>	<b>p&lt;.01</b>
Start-point	<b>p&lt;.05</b>	$p > .1$



## Issues to be adressed

- Only entering all the data (both Front and Back  $V_1$ s) into the comparisons does lead to such observation;
- Decomposing the data into respectively Front and Back  $V_1$ s produces seemingly uninterpretable results (possibly due to power / variability, use of t-tests not controlling for some sources of variability. . . );
- Normalize formant frequencies between speakers and / or vowels so we can compare Front and Back  $V_1$ s and speakers more satisfactorily within the statistical analyses (e.g. Gerstman, 1968);
- Apply this approach to regular cases of VH (Turkish, Hungarian. . . );

