

# VOWEL QUALITY DECAY AND PERCEPTUAL ASYMMETRIES IN LIGHT OF THE NEUTRALISATION HYPOTHESIS

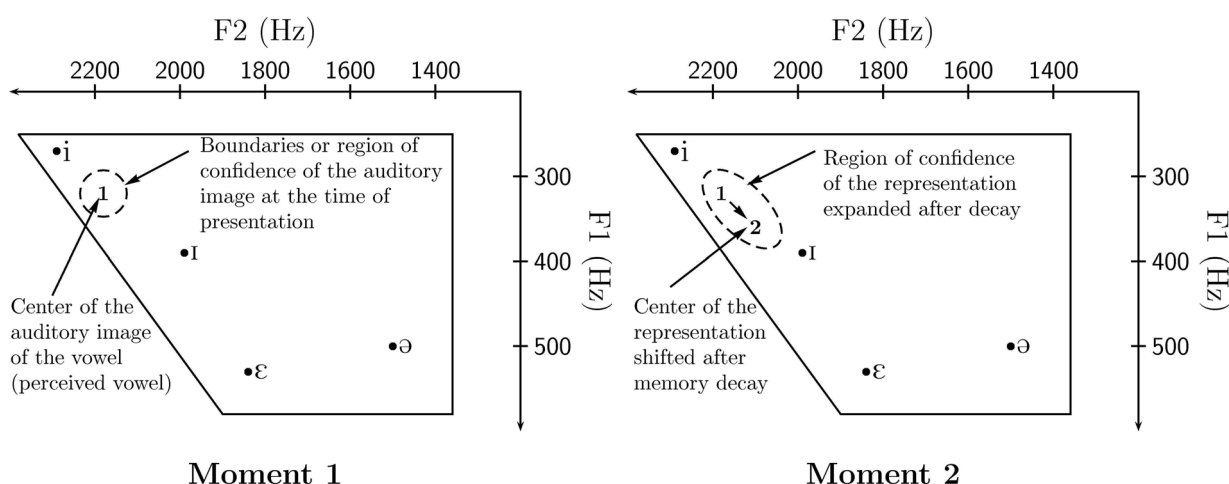
Charalampos Karypidis and Antonia Colazo-Simon  
*LPP - UMR 7018, Université de la Sorbonne Nouvelle - Paris 3 / C.N.R.S.*  
*ch.karypidis@gmail.com, simonantonia@hotmail.com*

## Abstract

*The present paper consists in replicating Repp and Crowder's (1990) neutralisation hypothesis according to which the auditory trace left by the first vowel in a pair decays toward a central point in the acoustic space, triggering order effects in perception. Two 'same-different' tasks were undertaken by French listeners. A prototype and four satellite tokens were synthesized for ten vowel categories of French; satellites were paired with the prototype in both orders of presentation. The focus of the two tasks undertaken was the possibility of [ə] being a target of decay for auditory traces on the F1/F2 and F2/F3 acoustic space. Results from both tasks failed to validate the initial hypothesis.*

In his seminal work *Elemente der Psychophysik*, Fechner (1860) introduces the concept of 'time-order error' which consists in the fact that subjects often over- or under-estimate the magnitude of a stimulus's properties depending on its order of presentation in a sequence of tokens. Albeit the growing literature on this topic, the mechanisms behind it remain obscure.

Using an /i/-/ɪ/ continuum, Cowan and Morse (1986) claim that discrimination was easier when the second vowel in a pair was closer to the /i/ endpoint. They suggest that the asymmetry in the perceived distance between two stimuli is triggered by a change in quality toward the centre of the vowel space corresponding to the region occupied by the vowel /ə/. The mechanism of this change in quality is presented in Figure 1.



**Figure 1. The vowel shift model, adapted from Cowan and Morse (1986).**

On the left is illustrated the auditory trace that the stimulus has left in memory 250 ms after its presentation. The dashed line indicates the vowel's 'confidence region'. In the middle of this region lies the centre of the mental representation that the listener has created for the stimulus. At moment 2, 2000 ms after the presentation of the stimulus, the confidence

region expands and the centre of the representation is displaced toward that of the acoustic space. The direction of this expansion depends on the location of the stimulus in the space. At moment 1, the stimulus lies close to the boundaries of the vowel space and its confidence region cannot expand toward a more extreme region; it thus expands toward the interior. In the case of a more interior, /ɪ/-like vowel, no such (acoustic) constraints are imposed and the expansion may occur in any direction.

By constructing an *ad hoc* protocol, Repp and Crowder (1990) assay this theory of decay which they call the ‘neutralisation hypothesis’. In the main experiment of their paper, the stimulus set was comprised by nine ‘prototype’ stimuli whose F1 and F2 values corresponded to the mean formant values of the vowels [i ɪ e æ a ɔ u ʊ] uttered by male American English speakers (Peterson and Barney, 1952). The vowel [ə], not included into Peterson and Barney’s study, was added to the stimulus set.

Our work consists in replicating Repp and Crowder’s study of the F1/F2 vowel space for French vowels. We have also expanded the hypothesis to the F2/F3 space. This paper is a continuation of Karypidis (2007) and Karypidis (2010) where results on the F1/F2 experiment were presented.

## Method

### Subjects

Twenty-three native speakers of French participated in each of the three subtasks of Experiment 1 (F1/F2) and twenty-one in the subtasks of Experiment 2 (F2/F3).

### Stimuli

Nine monophthongal vowels, [i e ε u o ɔ a y ø], corresponding to the average values for French vowels uttered by adult male speakers (Calliope, 1989), were synthesised with a cascade formant synthesiser (Klatt, 1980). The neutral vowel [ə] was added to the original set of nine vowels. Its properties were assumed to be those of a 17-cm uniform tube (F1 = 500, F2 = 1500, F3 = 2500, F4 = 3500 Hz). Raw values in Hertz for all ten vowels were converted into Bark scale with Traunmüller’s (1990) Hertz-to-Bark formula.

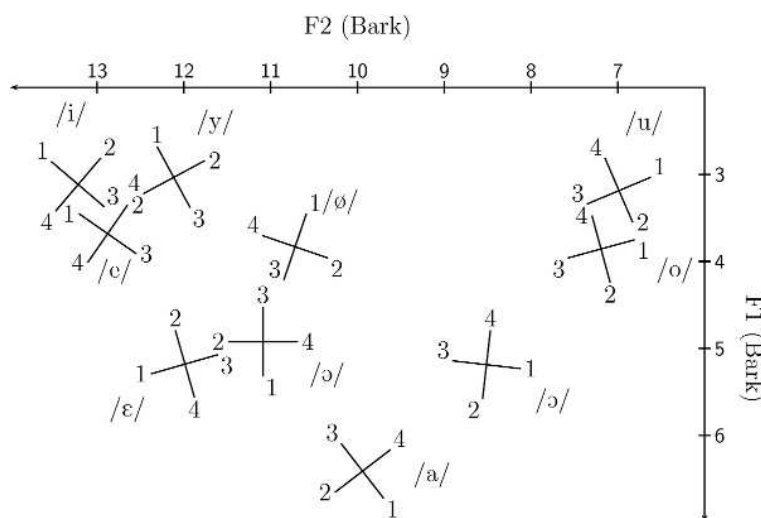


Figure 2. The fifty synthesised stimuli used in the F1/F2 task.

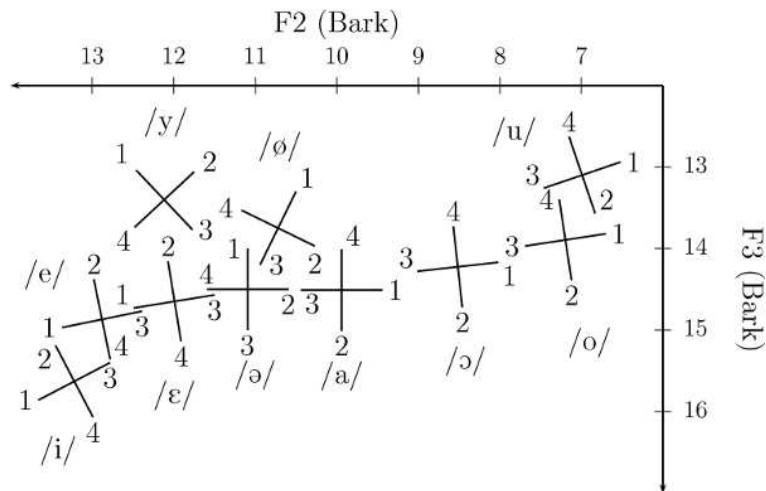


Figure 3. The fifty synthesised stimuli used in the F2/F3 task.

Each of the ten original prototypes (P) was surrounded by four satellites (S1–S4) in the form of a cross. S1–S4 were positioned on the endpoints of each cross, one axis of which pointed toward [ə]. Each arm was equivalent to a Euclidean distance of 0.4 Bark in Experiment 1 and 0.5 Bark in Experiment 2 between the prototype (located at the centre of the cross) and each satellite. S1 was located on the axis pointing toward [ə] and was positioned the furthest away from it; S3 was positioned the closest to [ə]. The rest of the tokens were numbered in a clockwise fashion. S1–S4 for /ə/ were arbitrarily numbered, and the two axes were parallel to the F1/F2 coordinates. A graphic representation of the two stimulus sets can be found in Figures 2 and 3.

### Procedure

Following Repp and Crowder, we chose a ‘same-different’ paradigm. For each phonetic category, tokens P and S1–S4 were paired with themselves (P/P, S<sub>i</sub>/S<sub>i</sub>). The four neighbours were also paired with the prototype in both orders (P/S<sub>i</sub>, S<sub>i</sub>/P). The intervals were set at 500 ms between the two members of a pair (interstimulus interval), at 1 second between pairs (intertrial interval) and at 3 seconds after each group of 13 pairs.

Participants were tested individually and tokens were presented binaurally over high-quality headphones which were previously calibrated at 70 dB SPL using a 1000-Hz pure tone. FLXLab 2.0 served as the interface. Six experimental blocks, each containing all pairs in random order (but identical for all listeners), were prepared. The first block was considered a training session and no feedback was offered. The objective was to familiarise listeners with the synthetic stimuli and the experimental task. Listeners were requested to judge whether the paired stimuli were absolutely identical or even slightly different by typing ‘d’ for ‘different’ (fr. *différents*) or ‘m’ for ‘same’ (fr. *mêmes*) on their keyboard. No feedback was given after each answer. Each stimulus set was broken down in three tasks which contained one of the following groups of vowels: /i e ɛ/, /y ø ə/ or /u o ɔ a/.

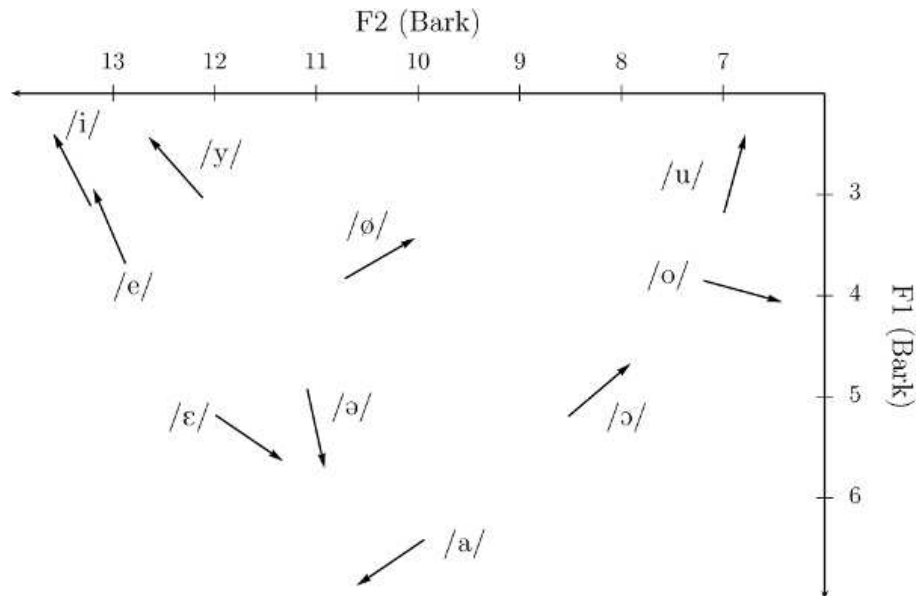
### Results and Discussion

Three main predictions are allowed within the neutralisation framework:

- [Pr1a] Positive order effects for S1: (P/S1)-(S1/P) > 0
- [Pr1b] Negative order effects for S3: (P/S3)-(S3/P) < 0

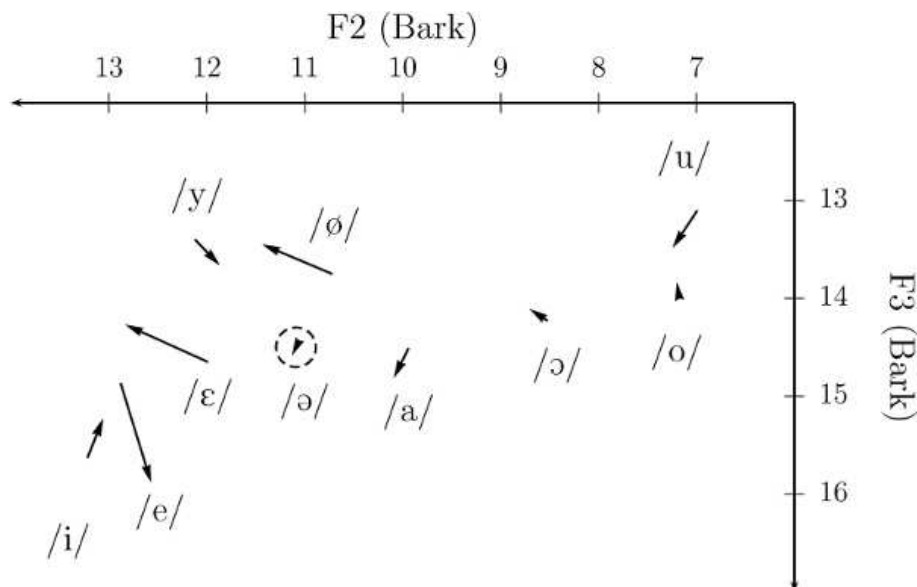
- [Pr2] No or minor effects for S2 nor S4

As regards Experiment 1, analyses of variance revealed a significant effect of Order for S1. An absence of order effects for S3 and significant order effects for S2 and S4 invalidated Pr1b and Pr2. In Experiment 2, the ANOVAs validated Pr1a and Pr2 but not Pr1a. Due to length constraints, we will not present a detailed version of the results.



**Figure 4. Vector representation of the order effects in Experiment 1.**

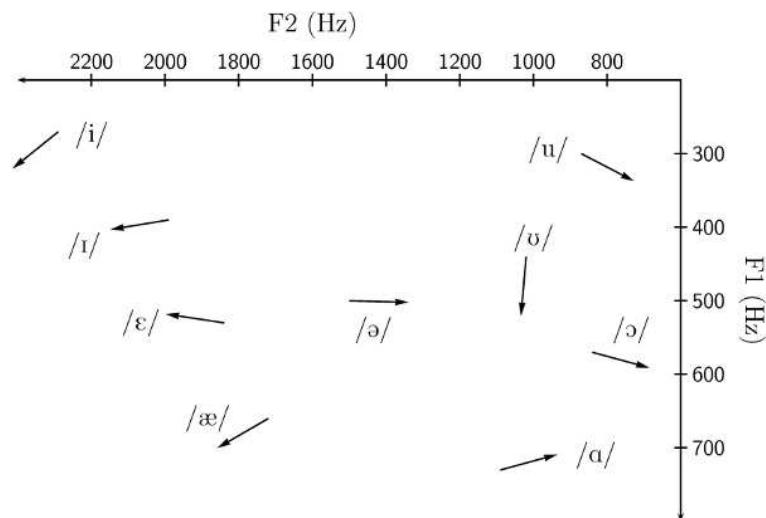
In order to discern a general tendency within each category, data for the four satellites was transformed into vectors with a vector calculator.



**Figure 5. Vector representation of the order effects in Experiment 2.**

The results of this method are available in Figures 4 and 5, where arrows indicate the order in which discrimination was easier. The neutralisation hypothesis predicts that arrows will point away from [ə]. Figure 4 reveals that [ə] was not the target of auditory decay in the F1/F2 plane. It appears that the target of front vowels was located around [ε] whilst the target of the nonperipheral vowels was situated around the middle of that subset.

In Figure 5 there seems to be no cohesion either within each of the three groups of vowels or for the entire set in the F2/F3 space. It is therefore not possible to detect a single target of auditory decay.



**Figure 6. Vector representation of the order effects found in Repp and Crowder (1990).**

We have subsequently transformed Repp and Crowder's data into vectors in order to compare it to our results. Figure 6 shows that vowels do not seem to decay toward a specific region in the F1/F2 acoustic space.

### Conclusion

Although the neutralisation hypothesis does not seem a viable model for explaining order effects in perception, this body of data should not discourage us from trying to understand the cycle of life of auditory traces in the precategorical and short-term memory. With modern cognitive or psycholinguistic theoretical frameworks regarding words as whole entities and not as concatenations of sounds which have to be identified one after the other and then combined in larger units, the temporal and quality aspects of auditory decay deserve more attention. This kind of work can provide answers to a question that has yet to be answered: are phonotactic constraints imposed by languages linked to the stability of the auditory traces of segments? That is, do languages juxtapose segments according to their auditory robustness in order to alleviate listeners' decoding mechanisms and prevent some sounds from getting drowned in a sequence?

### Acknowledgements

The authors would like to thank all listeners for their time and effort.

### References

- Calliope. (1989). *La parole et son traitement automatique*. Paris: Masson.
- Cowan, N., & Morse, P. (1986). The use of auditory and phonetic memory in vowel discrimination. *The Journal of the Acoustical Society of America*, 79, 500-507.
- Fechner, G. T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf und Hèartel.
- Karypidis, C. (2007). Order effects and vowel decay in short-term memory: the neutralization hypothesis. In *Proceedings of the XVI International Congress of Phonetic Sciences* (pp. 657-660). Saarbrücken, Germany.

- Karypidis, C. (2010). *Asymétries en perception et traitement de bas niveau: traces auditives, mémoire à court terme et représentations mentales (Asymmetries in perception and low-level processing: auditory traces, short-term memory and mental representations)* (PhD thesis). Université Paris 3 – Sorbonne Nouvelle, Paris, France.
- Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, 67, 737-793.
- Peterson, G. E., & Barney, H. L. (1952). Control methods used in the study of the vowels. *Journal of the Acoustical Society of America*, 24, 175-184.
- Repp, B. H., & Crowder, R. G. (1990). Stimulus order effects in vowel discrimination. *Journal of the Acoustical Society of America*, 88(5), 2080–2090.
- Trautmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America*, 88(1), 97–100.