Order effects and vowel decay in short-term memory: the neutralization hypothesis

Charalampos Karypidis

ORDER EFFECTS AND VOWEL DECAY IN SHORT-TERM MEMORY: THE NEUTRALIZATION HYPOTHESIS

Charalampos Karypidis

LPP - UMR 7018, CNRS / Université de la Sorbonne Nouvelle - Paris 3
ch_karypidis@yahoo.com

ABSTRACT

Information stored in short-term memory decays extremely fast compared to the content of long-term memory. The nature of this memory loss being little known (random or systematic), this paper examines the presentation order effect in the light of the neutralization hypothesis, according to which the first vowel in a pair decays, while retained in short-term memory, toward [ə]. Twelve French listeners participated in three AX roving discrimination sessions. For each of the 10 vowel categories of French, a prototype and four satellite tokens were synthesized. Tokens were paired with themselves as well as with the prototype (in both orders of presentation). Results revealed minor or major order effects within every phonetic category. However, the neutralization hypothesis could not account for at least half of these asymmetries. An alternative model, which adopts the periphery of the vowel space as a reference area, appears to fit best with our data.

Keywords: order effect, vowel perception, neutralization, short-term memory, peripherality.
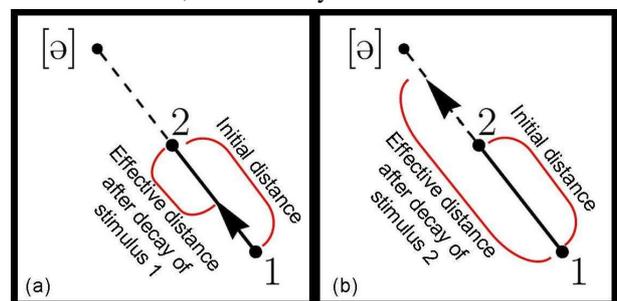
1. INTRODUCTION

The 'presentation order effect' (POE) - also known as 'phenomenon of asymmetry' - can be summarized in Polka and Bohn's (6) definition: "Asymmetries in vowel perception occur such that discrimination of a vowel change presented in one direction is easier compared to the same change presented in the reverse direction".

The order effect has been attributed to various factors such as peripherality (see Polka (6) for F1/F2; Best (1) for F1/F2/F3), focalization (9) and typicality (5). Using Rosch's (8) terminology, we will consider the aforementioned factors as cognitive reference points.

Cowan and Morse (3), while interpreting their results, suggested that in a pair of vowels V1V2, it is possible that the auditory trace left in memory by V1 (the first vowel presented in the pair) might gradually degrade toward a neutral (central) position in the vowel space which acts as a perceptual anchor. The neutral vowel would thus behave as a perceptual

Figure 1: Direction of decay predicted by the neutralization hypothesis for stimuli 1 and 2 in the order *1-2* (a) and in the order *2-1* (b). In (a), 1 moves toward [ə] and, at the same time, toward stimulus 2; in (b), stimulus 2 moves toward [ə] but this time, further away from 1.

magnet toward which a vowel is drawn whilst stored in short-term memory, waiting to be used or to be discarded in order to make room for the next chunk of information to be stacked away.

Figure 1 depicts the direction of decay of stimuli 1 and 2 presented in both orders (1-2 and 2-1). According to the neutralization hypothesis, when the tokens are presented in the order 1-2, stimulus 1 will degrade toward [ə] and, at the same time, toward 2, which is located on the axis of decay. Therefore, the initial acoustic difference will become smaller. In the order 2-1, on the other hand, stimulus 2 will move toward [ə] but this time further away from 1, rendering the initial difference greater. Therefore, discrimination scores in the order 2-1 will be greater than those for the order 1-2. More generally, **discrimination would be easier when in a pair, the first stimulus presented is closer to the adopted reference point**. Repp and Crowder (7), after constructing an innovative set of stimuli to examine this hypothesis (which they called the 'neutralization hypothesis', henceforward NH), came to the conclusion that vowels tend rather to be drawn toward the center of a given stimulus set.

This paper examines the NH for French vowels. A set of 50 stimuli spanning the ten vowel categories in French allowed us to observe the direction in which vowels are easier to discriminate.

Table 1: Formant values of the ten vowel prototypes in Bark.

Vowel	F1	F2	F3	F4
i	3.11	13.22	15.63	16.49
e	3.70	12.90	14.90	16.40
ɛ	5.20	12.00	14.60	16.30
a	6.40	9.90	14.50	16.20
ɔ	5.20	8.50	14.20	16.20
o	3.90	7.20	13.90	16.20
u	3.20	7.00	13.10	15.90
y	3.00	12.10	13.40	16.00
ø	3.80	10.70	13.80	16.10
ə	4.90	11.10	14.50	16.70

2. EXPERIMENT

2.1. Participants

In total, 18 French-speaking listeners (range: 21-47 years; mean: 28 years) participated in this experiment. Not all participants underwent all three tasks, thus 12 listeners were finally obtained for each of the three session. All reported being native speakers and having no known speech or hearing disorders.
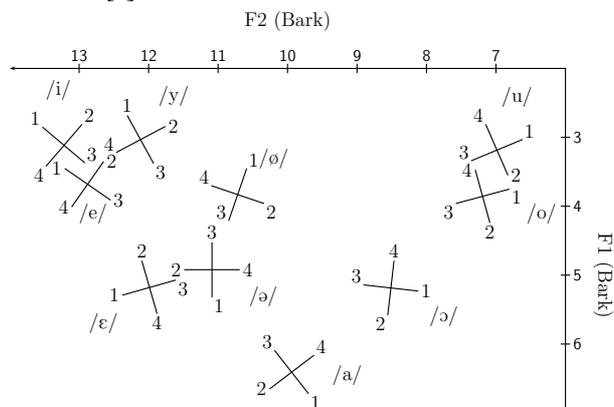
2.2. Stimuli

Nine monophthongal vowels, [i, e, ɛ, u, o, ɔ, a, y, ø], corresponding to the average values[1] of French vowels uttered by adult male speakers, were synthesized. The neutral vowel [ə] was then added to the original set of 9 vowels. Its properties were assumed to be identical to those of a uniform tube: $F1 = 500$, $F2 = 1500$, $F3 = 2500$, $F4 = 3500$ Hz. Raw values for all ten vowels were converted into Bark scale (11) in order to account for discrepancies in sensitivity of frequency perception (Table 1).

Each of the ten original prototypes[2] (P) was surrounded by four neighboring stimuli (N1–N4) in the form of a cross. N1–N4 were positioned on the endpoints of each cross, one axis of which pointed toward [ə]. Each arm was equivalent to a Euclidean distance (on the F1-F2 plane) of 0.4 Bark between the prototype and each neighbor. N1 was located on the axis pointing toward [ə] and was positioned the furthest away from it (contrary to N3 which was positionned the closest to [ə]). The rest of the tokens were numbered in a clockwise fashion. N1–N4 for the /ə/ category were arbitrarily numbered and the two axes were parallel to the F1-F2 coordinates. An F1-F2 plot of the fifty stimuli is found in Figure 2.

F3 and F4 were fixed independently within each category (the F3 and F4 values of each prototype were used for its four corresponding satellites). Duration of tokens was fixed at 250 ms and F0 contour

Figure 2: F1-F2 plot of the fifty synthesized stimuli. In the center of each cross lies the prototype. Neighbors (N) 1 and 3 lie on the axis pointing toward [ə].

was falling (100–80 Hz). A cascade formant synthesizer (4) was used for the preparation of the stimuli, which were matched in RMS energy at -10 dB using Sound Forge 6.0.

2.3. Procedure

For each phonetic category, tokens (P, N1–N4) were paired with themselves (P/P, N1/N1, ...) and the four neighbors were also paired with the prototype in both orders (P/N1, N1/P, ...). The Inter-Stimulus Interval (ISI) was fixed at 500 ms. Six experimental blocks, each containing all pairs in random order (identical for all listeners), were prepared. The first block was considered a training session.

Listeners were requested to judge whether the paired stimuli were absolutely identical or even slightly different by typing 'd' for 'différents' (eng. *different*) or 'm' for 'mêmes' (eng. *same*) on the keyboard. The first stimulus of the following pair was presented with a 1000-millisecond delay. No feedback was given after each answer.
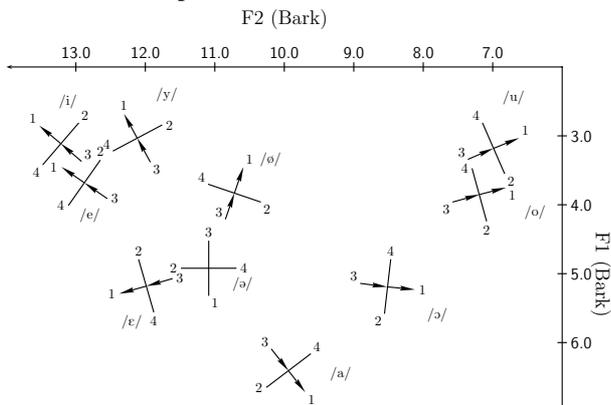
2.4. Results

According to NH, four main predictions (Pr) could be made (Figure 3):

- **Pr1**: positive order effect for N1 (P/N1>N1/P): discrimination scores for pairs in the order 'P/N1' would be greater than scores in the order 'N1/P';
- **Pr2**: negative order effect for N3 (P/N3<N3/P);
- **Pr3**: no particular order effect for N2 and N4;
- **Pr4**: no particular order effect for the [ə] category.

Two separate three-way (factors: Order, Neighbor, Vowel) ANOVAs were conducted, one on pairs involving N1 and N3 and another for N2 and N4

Figure 3: Order effects predicted by the neutralization hypothesis. Discrimination is easier when the stimulus presented first is located closer to [ə].

pairs. Scores for the [ə] category were excluded, because no order effects were expected.

In the case of the N1 and N3 pairs, a significant effect of Order [$F(1, 396) = 17.568$; $p < .001$] and Vowel [$F(8, 396) = 14.571$; $p < .001$] was found, as well as a Neighbor*Order interaction [$F(1, 396) = 23.295$; $p < .001$]. For N2 and N4 pairs, results revealed a significant effect of Order [$F(1, 396) = 19.525$; $p < .001$] and Vowel [$F(8, 396) = 8.311$; $p < .001$] as well as Vowel*Neighbor [$F(8, 396) = 11.379$; $p < .001$] and triple Vowel*Neighbor*Order [$F(8, 396) = 2.944$; $p < .005$] interactions. The effect of Order for N2 and N4 was evidently incompatible with NH. At the same time, the ANOVA on N1 and N3 could not inform us whether the direction of order effects was the one predicted by NH (cf. Figure 3). Therefore, a follow-up analysis (separate ANOVAs for each Neighbor) was conducted.

This time, a significant effect of Order was found for N1 [$F(1, 198) = 44.325$, $p < .001$], N2 [$F(1, 198) = 9.936$, $p < .005$] and N4 [$F(1, 198) = 9.627$, $p < .005$] but not for N3 [$F(1, 198) = .186$, $p = .667$]. The effect for N1 confirms Pr1. However, the order effect for N2 and N4, on one hand, and the absence of effect for N3, on the other, are incompatible with Pr3 and Pr2 respectively. The effect of Vowel was significant in all cases [$F(8, 198) = 10.888$; $F(8, 198) = 7.071$; $F(8, 198) = 5.783$; $F(8, 198) = 13.259$ for N1, N2, N3 and N4 respectively; $p < .001$ for all four neighbors].

2.5. Discussion

A better understanding of the asymmetries found in our data can be obtained with a graphical representation. In Figure 4, the arrows depict the order in which listeners discriminated considerably better.

We have chosen as a threshold of significant difference between two orders that proposed by Repp and Crowder, that is, 10%. We have later readjusted this threshold to 11%, given that some scores were critical (10.8%). Therefore, for a given pair (i.e. P/N1), an arrow pointing toward N1 suggests that discrimination was better in the order 'P/N1': (P/N1)-(N1/P)$> 11\%$]. As it was explained in the Introduction section, this is due to the fact that in this order, P is attracted toward [ə] and thus distances itself even more from N1.

NH predicts 18 order effects (Figure 3), two (N1 and N3) for each of the nine vowel categories of which only 11 are found in Figure 4. In addition, NH cannot account for the 14 additional asymmetries involving N2, N3 or N4. It is also worth noting that within the /ə/ category, order effects were found for 3 of four neighbors (Pr4). Practically all arrows exhibit a rather considerable consistency: almost all of them point toward the periphery (the edges) of the vowel space. This seems to agree with a series of papers on the role of peripherality in the asymmetry effect (see (6) for a review).
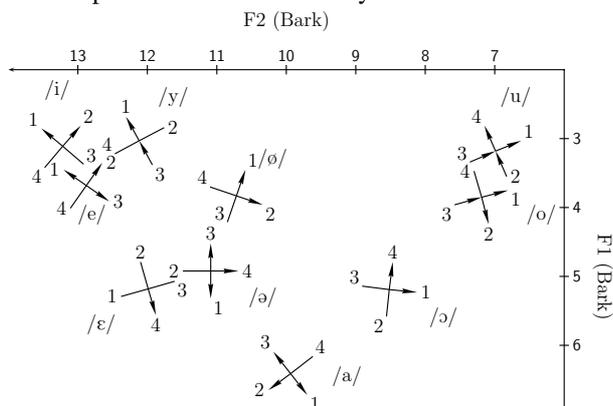
It appears thus that discrimination is easier when a peripheral vowel is presented second. In this case, one is to assume that the reference point is not a single point on the space but rather involves the whole perimeter (periphery) of the vowel space. However, if the asymmetry effect was due to the vowel decaying towards a reference point (the periphery), discrimination would be easier when the vowel closer to the reference point was presented first (see Figures 1 and 3). Arrows in Figure 4 exhibit an inversed polarity. This led us to the hypothesis that the presentation order effect is rather triggered by a contrast effect due to recency effects. Our postulate would thus be: "the more peripheral the second vowel, the stronger the constrast it generates with the first vowel".

Our data and our hypothesis on the role of peripherality are supported by Repp and Crowder's results, where the role of peripherality appears to account for more than 85% of the order effects. Moreover, the direction of the arrows found in Figure 4 is very close to that presented in Polka and Bohn's (6) Figure 1a, which offers an overview of the literature on infant data.

3. CONCLUSION

In this paper, we have examined the presentation order effect in the light of the neutralization hypothesis. Using a set of stimuli spanning the ten phonetic categories of French vowels and three AX roving discrimination tests, we have studied the extent to

Figure 4: Order effects for the 12 subjects. Arrows point to the direction in which vowel change was perceived more efficiently.

which the neutralization hypothesis can account for presentation order effects obtained throughout the entire vowel space. Results indicated that the aforementioned hypothesis does not fit well with our data. A graphical representation of the order effects revealed that peripherality appears to be a much more plausible predictor. A hypothesis was then stated that asymmetries are not due to vowel decay but that are rather triggered by the stimulus presented second in a pair. More precisely, in a pair of front vowels, the order effect would be triggered by the vowel that is more front (higher F2) via a retroactive contrast effect. The same tendency would be valid for high, low and back vowels.

Whether peripherality is a factor setting off asymmetries, a very important, two-fold question remains to be answered: which cognitive process is behind this phenomenon and whether the trigger is the first or the second vowel. NH suggests that the decisive factor is the direction of decay of V1 whilst the hypothesis of peripherality presupposes a retroactive contrast effect. Any attempt to answer these questions could have important theoretical implications in the field of speech perception, giving us an insight to how information is analyzed and stored in short-term memory.

Further experiments to address this question are currently being conducted. Only pairs containing the prototype and N1 of each category were included. In the first two tasks, the Inter-Stimulus Interval was equal to either 200 or 1000 millisecond. At a second stage, the three tasks described in this paper were undertaken by listeners with different linguistic background. Given that the periphery of the vowel space is, due to articulatory and acoustic contraints, similar to all listeners, order effects from both language groups (French and non-French)

are expected. For lack of space, the effect of experimental block and reaction times collected during our experiments will be exposed in a future paper.

5. REFERENCES

[1] Best, C. T., Faber, A. 2000. Developmental increase in infants' discrimination of nonnative vowels that adults assimilate to a single native vowel. *International Conference on Infant Studies* Brighton, 203–206.

[2] Calliope 1989. *La parole et son traitement automatique*. Paris: Masson.

[3] Cowan, N., Morse, P. A. 1986. The use of auditory and phonetic memory in vowel discrimination. *Journal of the Acoustical Society of America* 79.2, 500–507.

[4] Klatt, D. H. 1980. Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America* 67.3, 737–793.

[5] Kuhl, P. K. 1991. Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics* 50.2, 93–107.

[6] Polka, L., Bohn, O-S. 2003. Asymmetries in vowel perception. *Speech Communication* 41.1, 221–231.

[7] Repp, B. H., Crowder, R. G. 1990. Stimulus order effects in vowel discrimination. *Journal of the Acoustical Society of America* 88.5, 2080–2090.

[8] Rosch, E. 1975. Cognitive reference points. *Cognitive Psychology* 7.4, 532–547.

[9] Schwartz, J. L., Abry, C., Boë, L.- J., Ménard, L., Vallée, N. 2005. Asymmetries in vowel perception, in the context of the Dispersion-Focalisation Theory. *Speech Communication* 45.4, 425–434.

[10] Schwartz, J.- L., Escudier, P. 1989. A strong evidence for the existence of a large-scale integrated spectral representation in vowel perception. *Speech Communication* 8.3, 235–259.

[11] Traunmüller, H. 1990. Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America* 88.1, 97–100.

[1] In Calliope (2), [e, o, u, y, ø] were uttered in a pV context and [i, ɛ, a, ɔ] in a pVʀ context. Nonetheless, it is the only reference, of which we are aware, that comprises F3 and F4 values. Formant values obtained from preliminary recordings of French vowels in isolated context concur with the aforementioned values.

[2] We use the term 'prototype' here only for practical reasons. We do not presume that all listeners necessarily share the same category exemplars.