The perceptual asymmetry (PA) effect

Stålhammar, J.

journal: STL-QPSR
volume: 22
number: 2-3
year: 1981
pages: 001-007

http://www.speech.kth.se/qpsr
I. PERCEPTION

A. THE PERCEPTUAL ASYMMETRY (PA) EFFECT

J.U.J. Stålhammar

Abstract

Two case studies of what will be referred to as the Perceptual Asymmetry (PA) Effect will be introduced: 1) the nolla-hallon effect in hearing and 2) the perceptual focusing effect in vision. The former case is a spin-off from an evaluation of digit material for Automatic Speech Recognition - ASR - wherein the Swedish digit 'nolla' (zero) is transformed into 'hallon' (raspberry) in reversed playback. The weak aspiratory noise segment terminating the 'nolla' spectrum (the physiological explanation of which is a gradual glottal abduction gesture) is perceptually neglected in normal playback, but is perceptually detected in reversed playback. Two main hypotheses are advanced: (1) the forward masking hypothesis and (2) the linguistic structure dominance hypothesis, (data prediction). The experiments conducted advocate, with some reservation, hypothesis No 2. -- The latter case is likewise a spin-off, this time from research in Perception Physiology and Art Theory. In a target-space relation the target object is acting as a masker with fluctuating influence on the percepts of the surrounding space. The experiment shows that it is not possible to predict the interpretation of a given optical signal on the basis of a physical analysis of the spectral components alone. The two case studies indicate a functional similarity in the two senses, and focus upon the intricate relation between feature detection and feature decoding.

Introduction

During the evaluation of digit material for Automatic Speech Recognition, ASR, to be presented in Stålhammar (1982), previously in Stålhammar (1967, 1971), and in Stålhammar and Karlsson (1972), all of a sudden the present author discovered a raspberry in the digit basket. A raspberry in the sense of the Swedish digit 'nolla' (zero) -- simply played in reverse as shown in Fig. I-A-1, left. The signal was perceived as 'hallon' (raspberry) -- beginning with a laryngeal fricative /h/. The effect indicates a case of perceptual asymmetry as evidenced by the transcriptions underneath the two spectra. This effect was first reported by Stålhammar as the 'nolla'-'hallon' effect, Stålhammar (1967).
A similar observation in vision, 'the perceptual focusing' effect, is also demonstrated. This effect is a spin-off from research in picture decoding and synthesis, Stålhammar (1980, 1981). The experiments conducted indicate a functional similarity in the two senses, and focus upon the intricate relation between feature detection and feature decoding.

**Evaluation and Comments 1**

**The Nolla-Hallon Effect.**

From a physical point of view, the 'nolla' speech wave is carrying the so-called grave accent associated with an approximately symmetrical rising-falling FO contour, see Fig. I-A-1, right. This symmetry retains the bi-syllabic grave accent character when played backwards.

![Diagram of voice contour and spectrogram](image)

*Fig. I-A-1.* Forward (right) and backward (left) registration of test unit 'nolla'. Broadband spectrogram, FO and LP, BP, HP filtrations. Time code spike interval 20 msec for synchronization.
The spectrogram to the right represents the original recording, and to the left the reversed version is shown. Under scrutiny, a weak aspirative noise is revealed, terminating the 'nolla' spectrum (noticeable especially in the BP 2 - 4 kHz channel), thus the broad IPA transcription /nɔl:a+h/, where +h indicates the presence of aspiration. This post-vocalic aspiration of the open syllable is physiologically explained as a gradual glottal abduction gesture similar to an [h]-sound. Apart from the terminating aspiratory noise a possible cue to the /h/-percept could be an increased first formant damping associated with glottal abduction, and in addition, a change in the excitation function. This hypothesis, however, remains to be tested. The degree of first formant damping, moreover, can be seen to be speaker dependent, (Stålhammar, 1971, 1982). For a reference regarding voice source parameter dynamics, see Fant (1980). All these factors are part of an acoustic description of the signal and does not explain why the signal was perceived as including the segment /h/ when the tape was played in reverse and why the very same segment perceptually was neglected in normal playback. One explanation advanced is that we are dealing with a reflection of forward masking, but a more likely speculation is that this perceptual asymmetry is due to the operation of top-down processes, i.e., the percept is guided by the structure of the linguistic code, (speech mode). In the Scandinavian language family and in numerous other languages, the glottal fricative phoneme /h/ does not occur in syllable-final or word final-position. (See, e.g., Holmberg and Gibson 1979). However, it does occur in initial position. In the Swedish lexicon, the sequence '<allon>' is a non-existing morpheme. These facts obviously influence the percept of the acoustic signal drastically.

In an attempt to study the effect of this syllable-final /h/- noise further, a two-step operation was undertaken. In the first step the final aspiratory segment was gated out, i.e., the phonetic construct /nɔl:a-h/, (where -h indicates the absence of aspiration), was created, yet the prevailing interpretation of the truncated sequence was 'hallon'- /hål:ɔn/ in backward replay. The strong influence of semantics is mirrored. In the second step the original 'nolla' speech wave was electronically gated into two syllables -- /nɔ/ and /l:a(h)/ respectively.
A perception test was next carried out. The stimuli consisted of the /1:a(h)/ fraction with either the aspirative segment attached to the /1:a/ unit or removed from it, /1:a+h/, where ± indicates the presence or absence of aspiration, see Fig. 1-A-1. Theoretically the reversed versions should generate either of two possible lexical units, <all>, /al:/, or <hall>, /hal:/.

Test subjects performed very accurately in the sense that they consistently labelled the stimuli as <all> and <hall> depending on the absence or presence of aspiration, respectively. Apparently the order of magnitude of the expiratory noise amplitude was sufficiently above detection threshold, and constituting for the present stimuli a strong cue to the /h/ impression. The perceptual asymmetry might thus in reality be a reflection of brain plasticity. Information that is perceived as being irrelevant relative to the actual message structure is suppressed. This seems to be in agreement with the visual experiments to be presented below. The results of the tests, as a consequence, advocate hypothesis No 2. In order to verify the generality of this effect, several similar tests were performed using non-Scandinavian subjects in addition, Stålhammar (1967, 1971). One of these tests included the international female name 'Anna'. Again the dominating percept for the reversed version of 'Anna' was, as could be expected from the above experiments, 'Henna', i.e., another female name. It could be noticed here that in English the very spelling of this name is, as a matter of fact, often 'Hannah' including a final /h/ graph. -- The 'Anna/Hanna' asymmetry has later been corroborated by Holmberg for Swedish and Heike for German as shown in a general survey by Lindblom (1979).

The outcome of the tests indicates that when there are two physically similar morphemes to choose from, the subject (logically) is forced to become more attentive to the actual physical structure of the signal. In the doubtful case the physical properties of the signal are subsequently focused upon. In this way the interacting influence of 'form-based' (bottom up) phonetics and 'substance-based' (top down) phonology was demonstrated and a case of perceptual asymmetry was established, Stålhammar (1967).
The concept of perceptual asymmetry in hearing outlined above is related to the 'perceptual focusing' effect in vision, Stålhammar op. cit. In a series of visual experiments a side-effect pertinent to the present discussion was found. The experiment could be summarized as follows. A synthetic visual compound consisting of 1) a target object and 2) a surrounding space was submitted to a visual perception test. The target object was given either A) a physiognomic (organic, figurative, non-abstract) load, denoted <+physio> in Fig. I-A-2, left, or B) a non-physiognomic (non-organic, non-figurative, abstract) load, denoted <-physio> in Fig. I-A-2, right.

Fig. 1-A-2. Highly schematized illustration of the influence of target object, T.O., modes on the percepts of environmental space. (-) in extra-target locations = polychromatic space perceived as monochromatic. (+) in extra-target locations = the same polychromatic space perceived as polychromatic.
The feature \(+\text{physio}\) thus denotes an element with immediate associative load (such as a representation of a face) and \(-\text{physio}\) denotes an element without any immediate associative load. The surrounding space was a polychromatic construct. When the stimulus compound was generated with the target object under the \(+\text{physio}\) condition, A), the surrounding space was recalled as being monochromatic whereas when in the \(-\text{physio}\) mode, B), it was recalled as being polychromatic, see Fig. I-A-2. The psychological centre of interest, the target object, (gravity), actually masked off properties of the total signal compound, i.e., the target object acted as a masker raising the detection threshold for the parameters constituting the environmental space under condition A) to a higher level as compared to condition B). Again perceptual asymmetry and brain plasticity are reflected. The extra-target space components are perceived as being of minor importance in case A) and therefore suppressed. Accordingly it is not possible to predict the interpretation of a given optical signal (the polychromatic space in the present test) on the basis of a physical analysis of the spectral components alone, as indeed it is not possible to predict a listener’s judgment of a given acoustic signal as evidenced by the ‘nolla/hallon’ asymmetry on the basis of a physical analysis of the acoustic speech wave parameters alone.

By logical sequence, pattern perception is a far more complex phenomenon than pattern detection. The latter requires an ability of feature detection, the former, in addition, a knowledge of the psychology of detection. And here, in point of fact, lies the frontier (presently at least) between Artificial Intelligence and Intelligence in Art.
References


