Spatial Analysis of Hand Positions in French Cued Speech (LfPC)

Brigitte Bigi

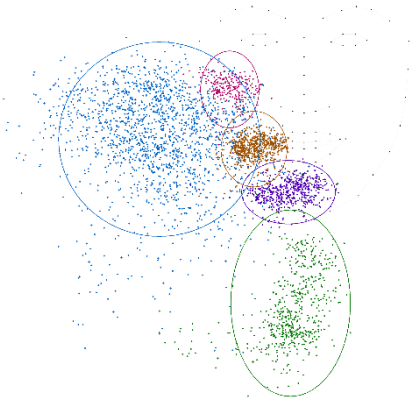
Speech production relies on both acoustic and visual cues. While lip movements convey valuable information, many phonemes remain visually ambiguous. Cued Speech is a visual coding system that complements lipreading by combining mouth shapes with hand cues—specific handshapes and positions around the face—that encode consonants and vowels. It provides full visual access to spoken language and supports phonological awareness, literacy development, and oral language acquisition in deaf or hard-of-hearing individuals. This study investigates the spatial distribution of vowel-related hand positions in French Cued Speech (LfPC), which uses eight handshapes for consonants and five distinct positions for vowels.

A corpus of LfPC was annotated at the phoneme level using SPPAS, then manually verified. Hand transitions were annotated frame-by-frame by a single expert. For each cued phoneme sequence, the frame at which the target position was reached was identified, allowing the extraction of spatial coordinates relative to the speaker's face. A total of 3,752 values were collected, normalized to a facial reference model, and analyzed to estimate mean positions and spatial variability.

Table 1 summarizes vowel-related positions, associated phonemes, number of cues, and the mean and standard deviation of x and y coordinates. Figure 1 illustrates these targeted positions.

Position	Phonemes	Number	Mean (x, y)	Stdev (x, y)
cheekbone	[ø], [ɛ̃]	230	(144, 397)	(70, 92)
chin	[ɛ], [ɔ], [u]	529	(425, 888)	(112, 76)
mouth	[i], [ã], [ɔ̃]	676	(260, 682)	(79, 91)
side	[a], [ə], [œ], [o], ∅	1750	(-193, 635)	(241, 233)
throat	[y], [e], [œ̃]	584	(435, 1419)	(143, 223)

Table 1 & Figure 1 : Distributional statistics and visual distribution



The four positions anchored on the face—cheekbone, chin, mouth, and throat—show relatively low variability, indicating stable production. In contrast, the 'side' position, which lacks a physical anchor point, displays the highest dispersion.

Nevertheless, mean coordinates are clearly differentiated across positions, confirming that *vowel cues are spatially distinct and consistently produced*. Further analysis revealed systematic variations depending on the nature of the task: natural reading versus pedagogical intent. This suggests that communicative purpose influences both precision and variability in cue production. Speaker-specific analysis confirmed the overall stability of the system, but also highlighted individual coding strategies, particularly for 'side' and 'throat' positions.

These results confirm the structured spatial organization of vowel cues in LfPC. Despite inter-speaker and task-related variation, the system ensures clear visual distinctions. The high dispersion of the 'side' cue emphasizes the role of anchoring in spatial accuracy. These findings inform linguistic modeling and practical applications, such as cue recognition or synthesis, and contribute to understanding the visual structure of Cued Speech.