SPEECH AND SIGN - IT'S ALL IN THE MOTION

Stina Ojala[1] *& Olli Aaltonen*[2]

[1]Department of Information Technology and [2]Department of Phonetics, University of Turku
{stina.ojala, olli.aaltonen}@utu.fi

ABSTRACT

Speech research has shown that vowels are less categorical than consonants, but a similar correlation in sign, i.e. between handshapes and place of articulation, is not yet known. The handshapes seem similar to vowels: they are continuum-like and follow coarticulatory principles. Here categorization and discrimination of handshapes were studied from the perspective of vowel perception. According to the results handshapes from the Finnish Sign Language handshape continuum transcribed as /G/-/X/ are perceived similarly than vowels varying systematically along a phonetic continuum. As in vowels, a phoneme boundary between signs can be found. In addition, there is a tendency for enhanced discrimination at the boundary zone. However, these results are typical to native signers only.

Keywords: Phonetics, sign language, articulation, mirror neuron system.

1. INTRODUCTION

The main element of the gestural/motor theory of speech perception, is that the main percepts of speech are not the acoustic patterns per se but the articulatory patterns that trigger the acoustic signal Liberman & Mattingly [7]. Recently, sign researchers have also promoted this idea, although in slightly different words, as suggested by Wilcox [18]: "making meaning with movement" - encoding of the meaning in articulatory gestures and trajectories. Both sign and speech are thus encoded in articulatory movements and then the only difference that remains is the channel. Both speech and sign obey coarticulatory effects and follow ease-of-articulation [15] and H&H-theory principles [8] with coarticulatory patterning [10].

All communication methods follow the parity principle: "what counts for a listener must count for a speaker and vice versa - otherwise there is no communication." [6]. The sensomotor integration together with the Mirror Neuron System (MNS) [13,14] can be seen as the main element of the language ability or, in other words, as the core of the sensomotor links between the observed articulatory gestures and the internal articulatory pattern memory traces. MNS presents "the missing link" of the language inside principle [9]: both sign and speech are manifestations of the same human language. MNS was first discovered with reaching for and grasping objects. Furthermore, mirror neurons were later discovered to trigger also in speech perception; the system was tuned from a general gesture and trajectory processing into more specific gestural patterns in speech [2]. Lately, recent findings in neurophysiological studies have shown that even signs are processed in Wernicke's area [11]. In speech processing, the so-called Broca–Wernicke complex can be seen as the core area within the MNS, the location where there are most connections, which in principle presents an analogue to monkey's brain regions with numerous mirror neurons.

The phoneme inventory of sign includes: handshape, place of articulation, orientation of palm and movement, and some studies regard mouthings as phonemes, as well, although their status in the phoneme inventory of sign is not yet clear. Here handshapes are studied as equivalent to vowels in speech. Handshapes were chosen because in sign research tradition they are the most thoroughly investigated group of sign phonemes. Handshapes are also used when classifying signs in Sign Language dictionaries that are arranged according to sign properties [17].

The purpose of this study was to investigate the identification and discrimination of the handshapes chosen from the handshape continuum generally transcribed in Stokoe notation [16] as /G/–/X/. The Stokoe notation is equivalent to IPA symbols in speech research. As IPA symbols, the symbols used in Stokoe notation refer to single elements in sign, which do not have a lexical meaning by themselves. Only certain combination of elements gathered together from the sign language phonemic inventory represents a lexically meaningful sign.

In visual terms the continuum reaches from a fist with protruding, fully extended index finger transcribed as /G/ to a fist with protruding, extremely crooked index finger, transcribed as /X/. In principle this continuum is equivalent to e.g. formant continua in vowels. This implies similar perceptual phenomena as in the categorization of vowels. For vowels, there are large individual differences in the locations of phoneme boundaries and category prototypes [1]. If a handshape continuum represents the same for signers as a vowel continuum represents for speakers, then similar effects in the perception of vowels could be expected.

2. MATERIAL AND METHODS

2.1. Subjects

There were 2 subjects from each of the 3 subject groups. The subject groups were: deaf, native FinSL signers, hearing FinSL interpreters, who have learned to sign in adult age and naïve speakers who had no command of any signed language. The subject groups were matched by age and sex. All subjects participated in the study as volunteers. The vision of all subjects was tested with clinical E-chart and it was within the normal range. The normal range is [1.25 – 2.00] which has been stated in DeValois & DeValois [4].

2.2. Study paradigm

The first two groups of the subjects participated in the production test set, which was done primarily to establish a basis for the synthetic stimuli design for the perception test set. The production test included three separate tasks: map task, ground plan explanation and free narration. This procedure elicited a sufficient number of handshapes from the desired continuum for the basis of the 3D synthesis for the perception tests. These three tasks were analogous to carrier words, carrier sentences and free narration in speech production studies.

The perception test was further divided into two separate tests: identification and discrimination tests. The AX paradigm was used when designing the discrimination test. In the identification test the subject was asked to label the stimuli as belonging to either category /G/ or /X/ and furthermore to rate the goodness of the stimulus with range [1-7] with 7 being the best possible and 1 recognizable. In the identification test given label, goodness-rating and reaction time (RT) were measured.

In the discrimination test the subjects were asked to state whether the stimuli in the seen stimulus pair were same or different (AX-paradigm). Subject's discrimination ability (d prime) was calculated along with RT measurements, which were automatically obtained in PXLab [12] presentation software.

2.3. Stimuli

The material for the 3D synthesis was gathered based on the native subjects' signed answers in the production tests. The synthetic continuum consisted of handshapes where the only protruding finger was the index finger while the other fingers made a fist. In one end of the continuum the protruding index finger was fully extended (/G/) and in the other end the index finger was as crooked as possible (/X/).

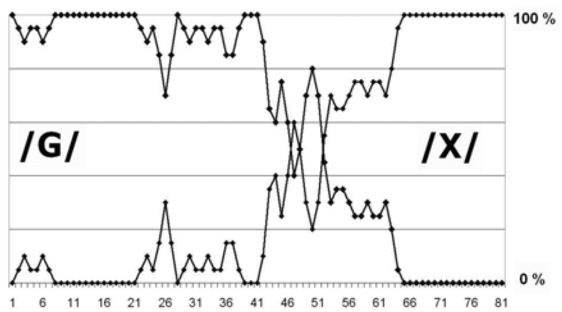
The 3D synthesis modeled the original handshapes produced in natural contexts by blue, oval shape figures. The blue color was used to avoid any conflicts with naturalness. The semisynthetic stimuli were used in the pilot tests, but they appeared to be a disturbing element for the concentration on the tasks in the experiments.

The synthetic stimuli were drawn on the continuum 0°–82° with 1° intervals. The degrees were measured from the tip of the index finger to its base. This measurement principle was possible because the muscles of the index finger are coordinated linearly, as stated in Darling et al. [3]. All the stimuli were included in the identification test, but only every other stimulus was chosen for the discrimination test, i.e. the stimuli were on 2° intervals. The intra-pair distance was thus 2° and the pairs covered the whole continuum 0-82°. This was done in order to reduce the concentration load for the participants of the study.

3. RESULTS

Identification and discrimination tests for these three different subject groups showed that groups differ in discrimination behavior but not in identification patterns. All subjects had individual differences in the place and steepness of the category boundaries. Also the naïve categorizers, i.e. those who were not previously used to handle visual input in identification tests were able to categorize between the different stimuli in the continuum.

Figure 1: Native signer's identification curves.



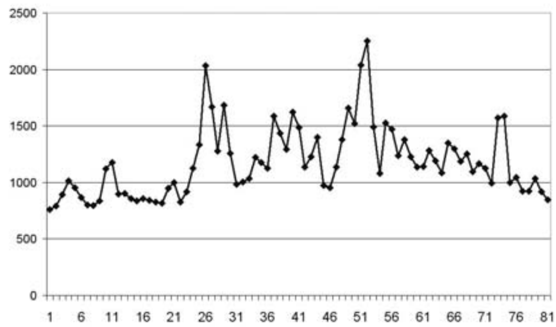
Both signers and naïve participants showed individual differences in the place and steepness of the category boundary (for example see Figure 1). The Sign Language interpreters, however, showed strikingly similar patterns in identification curves over the place and steepness of the category boundary within the group. Among all subjects the place of the category boundary varied from one individual to another, but the individual shapes of the identification function curves were stable over iteration of the identification tests.

Figure 2: Native signer's goodness-rating over the continuum

The goodness ratings were best within the category and poorer ratings were given to stimuli in boundary area (Figure 2). However, no clear prototype was found within a given category. This pattern was observed clearly in both signers and Sign Language interpreters. Also the naïve subjects could rate visual stimuli along the continuum.

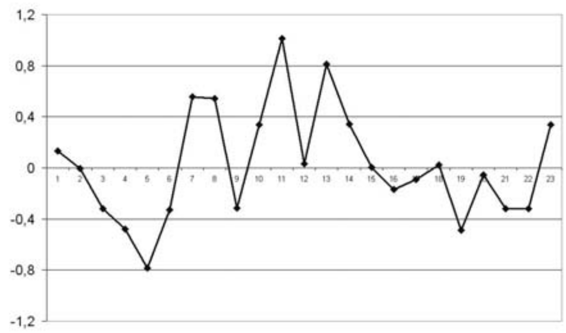
The reaction times were longer in the vicinity of boundary area than inside the boundary (Figure 3). The reaction times of the Sign Language interpreters also showed similar lenghthening effect in the vicinity of boundary area, but in naïve subjects' reaction time measurements similar effect was not present. The reaction times of the naïve participants were most delayed of all subjects, but there were no statistically significant differences between the reaction times of the Sign Language interpreter and the naïve participant groups [$p>0.5$]. The reaction times of the native signers were shorter than in the other two groups. This difference was considered statistically extremely significant [$p<0.0001$]. The statistical analyses were made with Wilcoxon matched pairs *t*-test. The statistical tests should be considered with regards to the small group sizes.

Figure 3: Native signer's reaction times over the continuum



The native signers had better discrimination ability in the vicinity of category boundary, but the discrimination patterns showed merely a continuum-like discrimination of the handshapes (Figure 4). The vertical axis in the figure represents d prime, the indicator of discrimination ability (the higher d prime, the better discrimination ability). The horizontal axis represents the discrimination ability over 4 degrees, i.e. two stimulus pairs.

Figure 4: Native signer's discrimination pattern over the continuum



The same effect was present in one of the Sign Language interpreters, but the other showed no increase of discrimination ability over the category boundary. The latter performance was similar to the discrimination ability pattern of the naïve participants who did not know sign. That is, the naïve participants were able to discriminate between different stimuli, but the ability remained at the same level throughout the continuum and no areas with enhanced discrimination ability were found along the continuum.

4. DISCUSSION

The identification patterns along the continuum show striking similarities to those presented by Aaltonen et al. [1] with vowel identification. Also the discrimination patterns coincide closely to those presented in speech research by various researchers. Thus, it would not be appropriate to regard the production and perception of speech and sign as two separate processes but one.

Furthermore, the findings underline the importance of articulation and articulatory movements in both the production and especially in the perception of sign as well as speech. Also, the results of the production test set give support to the ease-of-articulation principle suggested by Shariatmadari [16] which characterizes the articulatory gestures as the main substance of speech as well as sign. Similar coarticulatory phenomena are observable in signed narration as is in speech.

The patterns obtained from the native signers this far have been similar to those obtained from native speakers when categorizing their mother tongue. The absence of discrimination ability improvement in the naïve subjects could be a sign for the categorization of the continuum from a purely physical level, i.e. based on the physical differences of the stimuli. The performance of one of the Sign Language interpreters shows a native-like behavior when identifying and distinguishing handshapes. The other interpreter behaves like the naïve subjects when distinguishing between stimulus pairs, but the identification patterns with goodness-ratings are more similar to those of native signers and the other interpreter with better discrimination in category boundary area.

This framework also presents a testing ground for the plasticity of the human brain, as the visual processing seems to be more accurate and precise among the deaf than among hearing persons. This is revealed by the shorter reaction times in the results of the native signers than in other groups. The same precision and sensitivization of other senses when one is not present has also been found in the tactile precision of the blind, e.g. Kujala et al. [5].

5. REFERENCES

[1] Aaltonen, O., Eerola, O., Hellström, Å., Uusipaikka, E., Lang, A.H. 1997. Perceptual magnet effect in the light of behavioral and psychophysiological data. J. Acoust. Soc. Am. 101, 1090–1105.

[2] Arbib, M. 2002. The Mirror System, Imitation and the Evolution of Language. In: Nehaniv, C., Dautenhahn, K. (eds.) *Imitation in Animals and Artifact.* London: MIT Press.

[3] Darling, W.G., Cole, K.J., Miller, G.F. 1994. Coordination of index finger movements. *J. Biomechanics* 27, 479–491.

[4] DeValois, K.K., DeValois, R.L. 1990. *Spatial Vision*. Oxford Science Publications 14. New York: Oxford University Press.

[5] Kujala, T., Alho, K., Näätänen, R. 2000. Cross-modal reorganization of human cortical functions. *Trends Neurosci.* 23, 115–120.

[6] Liberman, A. 1993. Some assumptions about speech and how they changed. *Haskins Laboratories Status Report on Speech Research* 113, 1–32.

[7] Liberman, A.M., Studdert-Kennedy, I.G. 1985. The motor theory of speech perception revised. *Cognition* 21, 1–36.

[8] Lindblom, B. 1990. Explaining phonetic variation: A sketch of H & H theory. In: Hardcastle, W.J., Marchal, A.(eds.) *Speech production and speech modelling.* Dordrecht: Kluwer.

[9] López-García, A. (2005). *The grammar of genes: how the genetic code resembles the linguistic code.* Bern: P. Lang.

[10] Mauk, C. 2003. *Undershoot in Two Modalities: Evidence from Fast Speech and Fast Signing.* Academic dissertation. University of Texas at Austin.

[11] Nishimura, H., Hashikawa, K., Do, K., Iwaki, T., Watanabe, Y., Kusuoka, H., Nishimura, T., Kubo, T. 1999. Sign Language ’heard’ in the auditory cortex. *Nature*, 397, 116.

[12] PXLab software. http://www.uni-mannheim.de/fakul/ psycho/irtel/pxlab/ visited 6.3. 2007

[13] Rizzolatti, G., Arbib, M. 1998. Language within our grasp. *Trends Neurosci,* 21, 188–194.

[14] Rizzolatti, G., Craighero, L. 2004. The Mirror-Neuron System. *Annu. Rev. Neurosci* 27, 169–192.

[15] Shariatmadari, D. 2006. Sounds difficult? Why phonological theory needs ‘ease of articulation’ *SOAS Working Papers in Linguistics* 14, 207–226.

[16] Stokoe, W.C. 1960. *Sign language structure. An Outline of the visual communication system of the American deaf.* Studies in Linguistics. Occasional Papers 8. Department of Anthropology and Linguistics. University of Buffalo, New York.

[17] Suvi. Online dictionary of Finnish Sign Language. http://suvi.viittomat.net visited 6.3.2007.

[18] Wilcox, S. 1992. *Phonetics in Fingerspelling.* Studies in speech pathology and clinical linguistics 4. Amsterdam: John Benjamins.