2022. *Proc Ling Soc Amer* 7(1). 5166. https://doi.org/10.3765/plsa.v7i1.5166.

Facial locations in ASL based on production and perception data

Veronica Miatto*

Abstract. This study tests the phonological distinctiveness of eight facial locations in ASL, with both production and perception experiments. This kind of work is crucial because, due to scarceness of minimal pairs in sign languages, phonemic locations are difficult to determine. Moreover, claims made but not tested by previous theoretic models (Stokoe et al. 1960, Stokoe 1965, Battison et al. 1975, Battison 1978, Friedman 1977, Kegl & Wilbur 1976, Wilbur 1979, Sandler's 1989, Brentari's 1998) are here investigated, including whether locations that are predicted to be contrastive are indeed distinct (e.g. 'chin' vs 'mouth'). The specific goal of the first experiment (elicited production) is to determine what are the places of articulation, the aim of the second experiment (perception) is to determine if these places are contrastive.

Keywords. ASL; phonology; locations; categorical boundary

1. Introduction. Spoken and sign languages differ mainly on the modality in which communication happens; acoustic for spoken languages, and visual for sign languages. Brentari (2002) points out that the different modalities used in sign and spoken languages cause perceptual, articulatory, distributional, segmental, and lexical differences, which have consequences for the abstract phonological representations. One of the goals of linguistic studies is to account for both sign and spoken languages using the same abstract representations (Brentari 1998), since it is assumed that both modalities make use of the same underlying cognitive mechanisms.

One step towards reaching this goal is to advance our understanding of sign languages, since they are understudied, compared to spoken languages. This is the aim of this paper as well; more specifically, the present study seeks to contribute to determining which locations on the face are phonologically contrastive in American Sign Language (ASL). This topic has been debated since the inception of the subfield in the 1960s, with very different outcomes (Stokoe et al. 1960, Stokoe 1965, Battison et al. 1975, Battison 1978, Friedman 1977, Kegl & Wilbur 1976, Wilbur 1979, Sandler's 1989, Brentari's 1998), summarized in Section 2.1.

Here, I attempt to answer this question through an initial empirical study investigating the locations on the face where some common ASL signs are produced as well as testing the perceived discrimination of such locations, as well as some intermediate places on the face. To this end, first, an elicitation experiment was carried out, asking participant to indicate in which area of the face they would start signing a given sign. The areas investigated as possible places of articulation on the face include 'forehead', 'temple', 'cheek', 'chin', 'nose', 'ear', 'mouth', and 'mouth corner'. Next, a perception experiment investigated whether these areas are perceived as distinct through the use of nonce signs, and therefore can be considered distinct phonological categories[1].

* I would like to thank Tatiana Luchkina, Jenny Singleton and Mark Aronoff for the insightful comments and discussions during the development of this project. I would also like to thank the native signers that recorded the stimuli for the perception experiment, and all the participants of this study. It goes without saying that all errors are my own. Author: Veronica Miatto, Stony Brook University (veronica.miatto@stonybrook.edu).

[1] By phonological category I refer to "discrete elements that make up a phonological representation" (Boersma 2012: 207), which include phonemes, syllables, feature values and so on.

The paper is organized as follows: Section 2 will give background information on which locations have been considered contrastive in ASL and will summarize select studies that have dealt with phonemic discrimination in sign and spoken languages. Section 3 will present the elicitation experiment, and Section 4 will present the perception experiment. Finally, Section 5 will serve as conclusion.

2. Background.

2.1. CONTRASTIVE LOCATIONS ON THE HEAD IN ASL. The existent literature offers very different views with regards to which places of articulation should be considered contrastive in the sign domain. Stokoe et al. (1960) and Stokoe (1965), in their seminal work on sign structure, list many areas on the head as locations where a sign can be produced, including 'forehead', 'temple', 'cheek', 'ear', 'eyebrow', 'eyes', 'nose', 'lips', 'teeth', 'chin'. However, not all of them are considered contrastive. The four locations that Stokoe considers phonemic are: the upper face (including forehead and brow), the mid face (eye and nose region), the lower face (chin), and the side-face (temple, cheek, ear). What this approach to contractiveness implies is that 'mouth' and 'nose' are not considered contrastive, nor are 'temple', 'cheek', and 'ear'. However, 'temple' and 'forehead' are predicted be contrastive.

In the 1970s, much work was done to identify contrastive places of articulation. To illustrate, Battison et al. (1975) and Battison (1978) identified a number of additional sign locations in ASL (on the head and beyond). Specifically, they recognized 25 locations on the body (versus 12 for Stokoe). However, the list is not provided, so how many and which of these locations are on the head remains unspecified. Friedman (1977) expands on Stokoe's locations. She includes a very detailed list of locations on the head, for a total of 13. However, many of these locations are only used in special circumstances, such as for iconic signs (e.g. mouth and ear), or for double articulator signs (e.g. side of the forehead). Only four locations (forehead, center of nose, chin, and cheek) appear to be used in "unmarked" (i.e. non-special) contexts. It is also unclear if all these locations, or only the four "unmarked" ones are considered to be contrastive. Finally, Kegl & Wilbur (1976) and Wilbur (1979) distinguish signs in the neutral space and signs on the body with the feature [+/- contact]. In their work, they propose a complete feature matrix for location. On the head, they identify 9 locations that are distinguished using features such as [+/- bottom], [+/- lateral], etc. The 9 locations are: 'forehead', 'eyes', 'center nose', 'side nose', 'mouth', 'upper cheeks', 'cheeks', 'chin', 'ear'.

In Sandler's (1989) Hand Tier model of sign phonology, locations are a major phonological category. Instead of identifying them individually, she identifies a number of major locations, such as head, neck, shoulder, trunk, arm, and hand. She also identifies classes of features that distinguish the major locations. Some of these features are [contact]; [ipsilateral] and [contralateral], which identify signs on the same and opposite sides of the body as the dominant hand, respectively; and [high] and [low], which specify height within one of the major locations. Although a list of locations is not given, Sandler's model allows us to make predictions with regards to which locations should be contrastive in ASL: for example, only three different locations can be distinguished in the center of the face (center of the face uses the features [-ipsilateral] [-contralateral]): [+ high], [+low] and [-high] [-low]. Therefore, the four locations 'forehead', 'nose', 'mouth', and 'chin', which are all located in the center of the face, cannot all exist as locations in this model. Likewise, following Sandler's model, 'cheek' and 'temple' should not be contrastive, as there is no feature that could distinguish them.

More recently, in Brentari's (1998) framework, the place of articulation of a sign is on one of four major body areas ('major place') such as head, torso, body, and arm, and further falls under one of eight distinct locations ('distinction'). She predicts that while these areas and distinctions should be consistent cross-linguistically, their exact definition might not. For ASL, Brentari poses that, on the head, the eight places of articulation are 'top of the head', 'forehead', 'eye', 'cheek/nose', 'upper lip', 'mouth', 'chin', 'under the chin'.

While the aforementioned studies do offer some insights on contrastive places of articulation on the head in ASL, each account differs greatly from the others. Moreover, these predictions have never been tested experimentally.

2.2. PHONEMIC DISCRIMINATION IN ASL. To date, most studies on categorization and discrimination of sign language parameters have focused on handshape (Newport 1982, Emmorey et al. 2003, Baker et al. 2005, Best et al. 2010, Brentari & Eccarius 2011, Martinez 2015, among others) and found sharp categorization boundaries and a peak in discrimination performances around the category boundaries, meaning that signers are very good at discriminating pairs of handshapes across categorization boundaries, but not within categories. This is similar to what has been previously found in regard to consonant discrimination in spoken languages (see Liberman et al. 1957, Popper 1972, Pisoni & Tash 1974, among many others in support of this point, but also Carney et al. 1977 with incompatible results).

Newport (1982) investigated categorical perception of sign locations in addition to handshape in ASL. Her findings indicated that location-based contrasts (APPLE-ONION, for chin and temple) lack clear boundary peaks, similarly to the contrasts found for vowels in spoken languages (see for example Fry et al. 1962, Pisoni 1975, Tartter 1981, among many others). Similar results were found in Morford et al. (2008), for the contrast APPLE-ONION (chin-temple). Likewise, Emmorey et al. (2003) also looked at the location contrast between APPLE and ONION (chin-temple) and found a less sharp category boundary than that found for handshape, meaning that location boundaries are more diffused than those of handshapes, although they are still present. They also found no peaks in discrimination performances around the category boundary, meaning that speakers are very good at discriminating pairs of signs both within and across category boundaries. As they state, the handshape results are very similar to what has been found for consonants, while the locations results resemble those found for vowels discrimination in spoken language research.[2]

Notably, previous studies on location-based contrasts investigated only the contrast between the minimal pairs APPLE and ONION, located on 'chin' and 'temple', as far as the major area 'head' is concerned. The likely reason is that few minimal pairs exist in sign languages, let alone between contiguous locations on the head (van der Hulst & van der Kooij 2006, Eccarius 2008, Brentari & Eccarius 2011). An additional motivation to this study is that phonological contrasts on the head other than 'chin' and 'temple' have yet to be investigated in ASL. Due to the lack of minimal pairs, nonce signs will be used in the perception experiment.

3. Elicitation experiment. The first experiment is an elicited production experiment. The participants in the study were asked to indicate where they would start production of real signs by clicking with their mouse on a picture of a face. As the data does not come from direct measurements of participants' productions, the data collected only indirectly informs us on the locations of the tested ASL signs. The purpose of the experiment is to identify which areas on the face are

[2] Note that for other phenomena, such as syllabic structure, movement is usually considered vowel-like, while location is considered consonant-like (Corina 1990, Sandler 2012).

possibly contrastive in ASL by determining the areas of the face that are spatially distinct from one another in sign elicitation.

3.1. METHODS

3.1.1. PARTICIPANTS. 22 ASL signers were recruited for the elicited production experiment, which was administered online through Qualtrics (Qualtrics 2005). The mean age of the participants was 34.5 y.o. (st. dev. = 9 years). Seven identified as male and 15 as female. Seventeen participants were from the East coast (12 from the state of New York), and 5 from the Midwest. Nineteen participants were native or near native signers3, while 3 were high intermediate signers. Their proficiency was determined on the basis of a detailed questionnaire on their linguistic background and on two self-assessed proficiency tests. The data from all these participants was analyzed.

3.1.2. MATERIALS & PROCEDURE. For the elicited production task, 51 signs were selected, articulated in 8 areas on the face: 'cheek', 'chin', 'ear', 'forehead', 'mouth', 'mouth corner', 'nose', 'temple'. They were selected based on sign descriptions from the *American Sign Language Dictionary* (Sternberg 1998), and then cross-compared with two other online dictionaries to check for dialectal consistency: *handspeak.com* and *spreadthesign.com*. The experiment stimuli included between 7 and 9 signs for the areas 'cheek', 'chin', 'forehead', 'mouth', 'nose', and 'temple', and 3 signs for the areas 'mouth corner' and 'ear', which seem to be more marginal locations in ASL. In addition to being marginal, 'mouth corner' and 'ear' seem to be iconic locations too, as the most signs found and elicited are highly iconic (e.g. ADDICTED and DROOL for 'mouth corner', and EARRING and EAR for 'ear'). Twenty filler signs that are not articulated on the face were also elicited (e.g. RING, NAME, BREAD, and so on) for a total of 71 signs.

Participants were recruited through Mechanical Turk and through word of mouth, after an initial screening which required that they translate into English three ASL sentences presented in a video. In the experimental task, delivered in two blocks, participants were asked to point at the exact location where they would start to sign a word on a 600x600 picture of a forward-facing female face. They did so by clicking on the image of the face using their mouse or trackpad. They had the possibility of clicking outside of the face if they did not think that the sign was performed on the head. The output data resulting from each click was both visual, in the form of heatmaps, and numeric, in the form of coordinates (x, y). The latter had pixel as their unit of measure and used one of the corners of the image as the origin.

Additionally, a self-assessed language proficiency test was administered between the two blocks. It was adapted from the Sign Communication Proficiency Interview (SCPI) Rating Scale (Newell et al. 1983), and was used for self-rating following Stauffer (2011). The test consisted of can-do indicators describing what a signer should be able to do at each proficiency stage. Participants chose the indicator (among 7, with 7 being native proficiency) that most closely described their ASL proficiency. The mean score was 6.0 out of 7. A general linguistic questionnaire was also administered at the end (loosely adapted from Marian et al. 2007). Among other things, the questionnaire asked that the participants rate their ASL proficiency on a scale from 1 (beginner) to 10 (native) (mean score 9.2/10). The purpose of both the self-assessed test and questionnaires

[3] By near native signer I refer to signers, usually born Deaf or Hard of Hearing (HoH), who did not start learning ASL from birth. However, they usually started learning ASL before puberty, their proficiency is close to native and they use ASL on a daily basis.

was to ensure that participants were eligible for participation, as the study was conducted online and no interaction with the experimenter was possible.

3.1.3. ANALYSIS. For the purpose of this discussion, “data point” and “token” will refer to the pixel location where a participant selected the “start of the target sign”. Data points that were not on the face were eliminated, as well as the data that was completely off target. For example, one data point for ‘teach’ was eliminated because the eye was selected, instead of the temple which is the target location for that sign. Moreover, all data points on the right side of the face were collapsed with corresponding data points to the left side, to facilitate the statistical analyses. A total of 1042 out of 1122 tokens were analyzed.

The eight locations were divided into four ‘target groups’, made up of adjacent locations, so as to facilitate the statistical analysis. Target Group 1 compares the locations ‘chin’, ‘mouth’ and ‘nose’; Target Group 2 compares the locations ‘cheek’, ‘mouth’, and ‘mouth corner’; Target Group 3 compares the locations ‘cheek’, ‘forehead’ and ‘temple’; while Target Group 4 compares the locations ‘chin’, ‘cheek’ and ‘ear’.

A total of eight linear mixed-effects models were employed in the statistical analysis in R (R Core Team 2019). The dependent variables were either X or Y coordinates. The models that use the X values as dependent variable investigate whether locations differ on the horizontal axis (abscissa), while the models that use Y as dependent variable investigate whether locations differ on the vertical axis (ordinate). Therefore, the four groups of locations (see above) each had two models, one that looked at how much locations differ horizontally (X) and vertically (Y). The predictor of all models was location, with random slopes for participant and sign (in (1)):

(1) lmer (X ~ location + (1 | Participant) + (1 | word), data=table)

3.2. RESULTS. Estimates and t-values from the models are reported in Table 1 (significant effects are indicated with an asterisk).

Group 1	X	Y	Group 2	X	Y
-chin+mouth	Est=1.34 t=0.34	Est=14.42 t=3.42*	-cheek+mouthc	Est=32.32 t=6.85*	Est=-23.32 t=-2.37*
-chinmouth+nose	Est=3.41 t=1.03	Est=102.1 t=28.19*	-cheekmouthc+mouth	Est=55.85 t=16.56*	Est=-46.98 t=-6.67*

Group 3	X	Y	Group 4	X	Y
-temple+forehead	Est=39.31 t=5.89*	Est=28.15 t=2.27*	-cheek+chin	Est=70.86 t=17.59*	Est=-69.02 t=-12.33*
-templeforehead+cheek	Est=8.68 t=1.37	Est=173.87 t=14.51*	-cheekchin+ear	Est=-76.86 t=-16.37*	Est = 88.03 t=13.50*

Table 1. Results of the elicitation experiment with estimates and t-values

Target Group 1 compared the areas ‘mouth’, ‘chin’ and ‘nose’. Contrast setting for the factor ‘location’ was coded as follows: ‘chin’ = -1/2, -1/3; ‘mouth’ = +1/2, -1/3; ‘nose’ = 0, +2/3. Results show that the distance between the tested locations is not significant in terms of the X values (i.e. horizontally). However, distance between these locations is significant in terms of the Y values (i.e. vertically). Target Group 2 compared the areas ‘cheek’, ‘mouth’ and ‘mouth corner'. Contrast setting for the factor 'location' was coded as follows: 'cheek' = -1/2, -1/3; 'mouth' = 0, +2/3; 'mouth corner' = +1/2, -1/3. Results show that distance between locations is significant both in terms of the X and Y values (i.e. horizontally and vertically). Target Group 3 compared the areas 'cheek', 'forehead' and 'temple'. Contrast setting for the factor 'location' was coded as follows: 'cheek' = 0, -2/3; 'forehead' = +1/2, +1/3; 'temple' = -1/2, +1/3. Results show that distance between locations is significant in the X values (i.e. horizontally) for 'temple' vs 'forehead', but not between 'temple-forehead' and 'cheek'. However, distance between locations is significant in the Y values (i.e. vertically). Finally, Target Group 4 compared the areas 'cheek, 'chin' and 'ear'. Contrast setting for the factor 'location' was coded as follows: 'cheek' = -1/2, -1/3; 'chin' = +1/2, -1/3; 'ear' = 0, +2/3. Results show that distance between locations is significant both in the X and Y values (horizontally and vertically).

3.3. DISCUSSION. Based on the results of the elicited production task, all areas are statistically different from each other, either horizontally, vertically, or both. This suggests that the eight locations ('chin', 'cheek', 'nose', 'mouth', 'forehead', 'mouth corner', 'temple', 'ear') are at the very least phonetically distinct, and potentially contrastive at a phonological level as well.

Some other qualitative observations based on the heatmaps are of note. As a reminder, heatmaps for each sign were produced by Qualtrics using raw data from all participants. The significant difference between the areas 'mouth' and 'chin' was somewhat surprising, because of a considerable overlap between the two areas. This can be seen in Figures 1 and 2, which are the two output heatmaps for the signs RED and MISTAKE, respectively, and are exemplars of how participants responded for signs located on the mouth and chin.

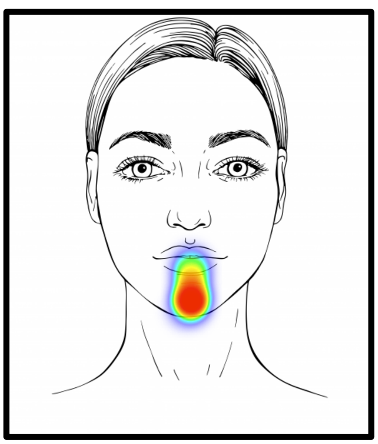

Figure 1. Heatmap output for the sign RED, located on the mouth

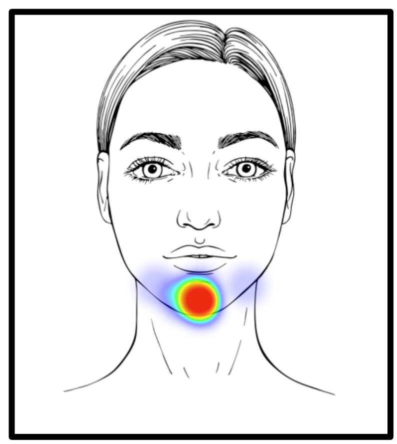

Figure 2. Heatmap output for the sign MISTAKE, located on the chin

Despite the visibly appreciable spatial overlap between signs located on the mouth and chin, the areas 'chin' and 'mouth' still look consistently different. More specifically, the area 'mouth' seems to include the center of the mouth and then lowers vertically to include the chin. This was very consistent across all signs, with the notable exception of EAT, which is produced exclusively on the mouth, as shown in Figure 3. On the other hand, the area 'chin' is located on the chin, and can spread towards the jaw, as seen in Figure 2.

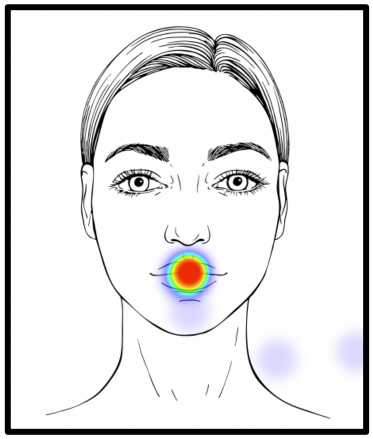

Figure 3. Heatmap output for the sign EAT, located on the mouth

Similarly, the areas ‘temple’ and ‘forehead’ overlap as well, although it seems that signs located on the ‘temple’ do not spread to the center of the forehead (Figure 4 for CONCEPT) as much as signs located on the forehead spread to the sides (Figure 5 for CLEVER).

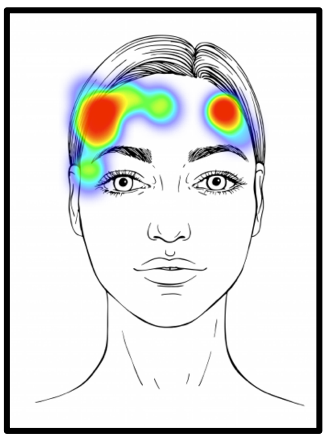

Figure 4. Heatmap output for the sign CONCEPT, located on the temple

Figure 5. Heatmap output for the sign CLEVER, located on the forehead

Moreover, while some areas occupy a large portion of the face, such as the forehead in Figure 5, some others occupy smaller regions of the face, indicating that the (potential) phonemic places of articulation are not equally large or evenly distributed. In these smaller areas, ‘nose’, ‘ear’, and ‘mouth corner’, we also do not see much overlap with other areas (except a bit of overlap between ‘mouth corner’ and ‘cheek’), or spreading. This is possibly due to the fact that these areas are physically small. Examples can be seen in Figures 6 (IGNORE), 7 (EARRING), and 8 (ADDICTED). Finally, the area ‘cheek’ appears to be quite distinct from others, with small overlaps with ‘mouth corner’ and ‘chin’, as can be seen in Figures 9 (APPLE) and 10 (GUM).

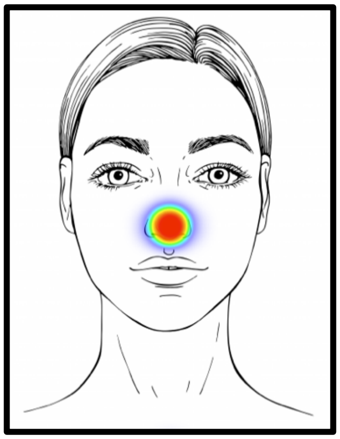

Figure 6. Heatmap output for the sign IGNORE, located on the nose

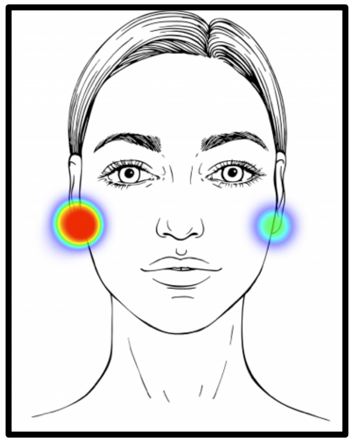

Figure 7. Heatmap output for the sign EARRING, located on the ear

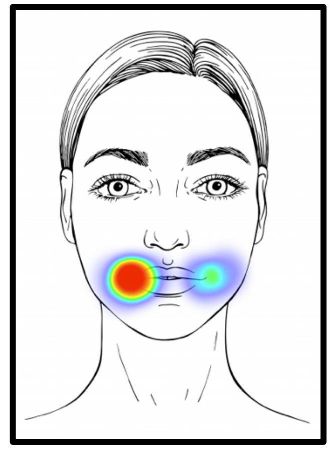

Figure 8. Heatmap output for the sign ADDICTED, located on the corner of the mouth

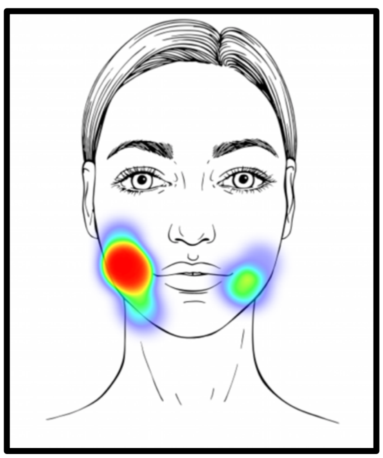

Figure 9. Heatmap output for the sign APPLE, located on the cheek

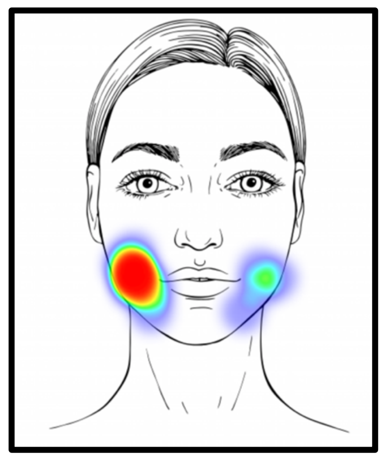

Figure 10. Heatmap output for the sign GUM, located on the cheek

In summary, all the facial areas here investigated ('forehead', 'temple', 'cheek', 'ear', 'mouth', 'mouth corner', 'nose', 'chin') appear to be distinct from each other. That said, it is important to remember that these first conclusions are completely based on indirect elicited production data. They would have to be confirmed with additional research, including a perception experiment. Moreover, all the signs elicited here were presented in isolation. It would be important to examine signs in co-articulation and context and how that affects place of articulation.

4. Elicitation experiment. The second experiment is a perception experiment, whose purpose is to identify which areas on the face found in the first experiment ('forehead', 'temple', 'cheek', 'ear', 'mouth', 'mouth corner', 'nose', 'chin') are perceived as phonologically contrastive in ASL.

Based on previous literature, I expect that signers will be generally very good at discriminating the nonce signs but will be worse when the contrast between two nonce sign locations is within a phonological boundary (Newport 1982, Emmorey et al. 2003, Morford et al. 2008).

4.1. METHODS

4.1.1. PARTICIPANTS. 8 ASL signers (mean age = 36.7, 7 females) were recruited for the second experiment. Seven participants were from the state of New York or New Jersey, while one was from California. Seven participants were native or near native signers, and one 1 was a high intermediate signer (although HoH). Their self-rated mean proficiency score is 9/10, and the mean score from the SCPI is 5.9/7.

4.1.2. MATERIALS. Since ASL lacks minimal pair signs for all possible pairwise locations included in this study, stimuli consisted of signs without a lexical meaning (nonce words). The stimuli were presented in the form of videoclips and were recorded by two native signers of ASL (they were not participants in the experiment). The model signers were instructed to make up signs using four chosen handshapes (D, H, bent V, X) in fourteen different locations (see Figure 11), which included the eight locations identified in Experiment 1, as well as in six additional locations that are intermediate between two major neighboring locations (for example, the intermediate point between 'chin' and 'cheek'). The selection of intermediate points was left to the judgements of the native signers, who could see themselves on a screen. The native signers recorded a total of 68 different signs, both as a word list, and within the frame sentence 'I sign […] now'.

4.1.3. PROCEDURE. The experiment took place partly on Qualtrics (Qualtrics 2005) and partly on the online version of Psychopy (Peirce 2007), PsychoJS, hosted on the website Pavlovia (www.pavlovia.org). Participants started on Qualtrics and were given the same linguistic background questionnaire as in the first experiment. They were then redirected to Pavlovia, where they completed the ABX task in two blocks. On each trial, participants were presented with two video clips of nonce signs (A and B) differing minimally by location. The minimal pairs represented one of the contrasts mentioned above. A third nonce sign (C), identical to either A or B, was then presented, and participants decided whether C was identical to A or B by pressing a designated key on their keyboard. They were allowed to watch the video clips only once.

Participants were exposed to a contrast once per block. Since there were two blocks, participants saw a contrast twice in the ABX task (x8 participants = 16 tokens per contrast). Both the signer and handshape used in a trial were randomly decided among those available (4 handshapes and 2 signers) during the design of the experiment. This means that the combination of handshape/contrast/signer in a trial was fixed in the experiment and the same for all participants. Both signer and handshape were identical in a singular trial. The combination of handshape/signer in a contrast was not controlled across blocks, so generally a location contrast was paired with a different handshape/signer combination across blocks.

Participants were then administered the SCPI self-assessed proficiency test (the same as in the first experiment), and a same/different task with the same nonce signs presented in carrier phrases. Participants indicated whether the sentences used the same signs by pressing designated keys on their keyboards. In this task as well, minimal pairs in each trial only differed by location, meaning that the handshape and signer assigned was random, but constant among participants. Participants were exposed to a contrast once per block, meaning that in total 16 tokens per contrast were collected.

The list of all contrasts between main adjacent areas investigated in the experiment is the following:

- Chin-mouth*
- Forehead-temple*
- Cheek-mouth corner*
- Mouth-mouth corner
- Chin-mouth corner
- Cheek-ear

- Nose-mouth corner
- Cheek-nose
- Mouth-nose*
- Cheek-mouth*
- Cheek-chin*
- Ear-temple
- Nose-forehead
- Nose-temple
- Cheek-temple

All investigated locations are shown in Figure 11. Some locations, for example ‘chin’ and ‘cheek’ (#12 and #5 in Figure 11, respectively), were not only contrasted with each other, but also with the intermediate point between the two (#14). Contrasts that have an intermediate point are indicated with * in the above list. The latter are also the contrasts investigated in the ABX task. The same/different task investigated the whole list.

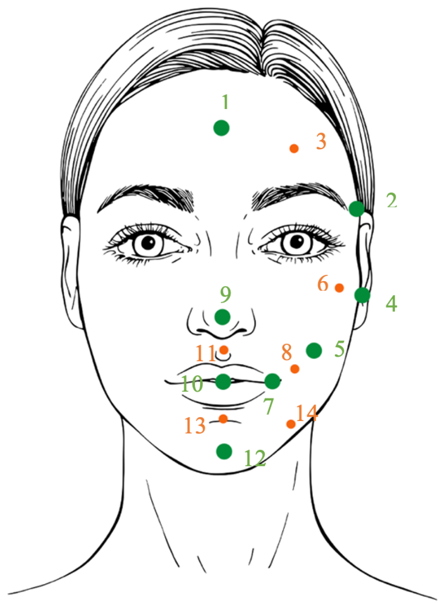

Figure 11. Location of main (green) and intermediate (orange) areas

4.1.4. ANALYSIS. In the ABX task, a total of 304 data points were collected, while for same/different task a total of 431 data points were collected. For the two tasks, both accuracy rates and reaction times were recorded and analyzed. Tokens yielding a reaction time that was more than 5000 milliseconds (ms) were removed from analysis (4 in the ABX task, and 5 in the same/different task). Mean reaction times per task were then computed, and tokens more than three standard deviations from the mean were deleted (9 in the ABX task, 14 in the same/different task).

Reaction times (dependent variable) were analyzed with linear mixed effects models, with location being the predictor and participant a random slope, as in (2). The binary responses correct/wrong (for the ABX task) and same/different (for the same/different task) were analyzed with generalized linear mixed effects models as dependent variables, with location being the predictor and participant a random slope, as in (3).

(2) lmer (reaction time ~ location + (1 | Participant), data=table)
(3) glmer (correct ~ location + (1 | Participant), family=binomial, data=table)

Due to the large number of contrasts, and the limited number of tokens per contrast, contrasts were divided into ‘main’, meaning the contrast between two main areas (those found in experiment 1), and ‘mid’, meaning a contrast between a main area and an intermediate one. This binary category is what is referred to with the predictor ‘location’. Contrasts for ‘location’ were coded as follows for all models: main = -0.5, mid = +0.5.

4.2. RESULTS. Table 2 reports accuracy rates and reaction times in ms per contrast (numbers in the first column refer to intermediate locations shown in Figure 11) for the ABX task. The table is ordered presenting a contrast between main areas (bolded), followed by the contrasts between those areas and their intermediate point (if present).

Contrast	Accuracy	Rt(ms)	N of tokens
Temple - Forehead	100%	436	15
Forehead - 3	80%	689	15
Temple - 3	100%	574	15
Mouth corner - Cheek	94%	457	16
Mouth corner - 8	80%	613	15
Cheek - 8	100%	428	14
Mouth – Mouth corner	100%	538	16
Chin - Cheek	100%	531	15
Chin - 14	87.5%	590	16
Cheek -14	100%	597	16
Mouth - Chin	100%	436	14
Mouth - 13	75%	533	16
Chin - 13	57%	577	15
Nose - Mouth	100%	453	15
Nose - 11	100%	526	16
Mouth - 11	100%	537	16
Cheek - Temple	94%	586	16
Cheek - 6	79%	692	14
Temple - 6	100%	613	16

Table 2. Accuracy rates and reactions times (ms) for the ABX task

Accuracy rates are fairly high, with several contrasts reaching 100% accuracy rates, and only one slightly above chance level, suggesting that the participants were very good at discriminating the stimuli. However, lower accuracy rates are almost always found in the contrasts between one of the major areas and one intermediate area, suggesting that these contrasts are harder to discriminate (e.g. the contrast mouth-chin has a higher accuracy rate than the contrasts mouth-13 and chin-13). This is illustrated in Figures 12 and 13, which reports mean reaction times and accuracy rates per type of contrast (main-main vs main-intermediate), respectively. This is also confirmed by the statistical analysis; contrasts between major places are more likely to be discriminated correctly than those between a major and an intermediate place (95% confidence interval [henceforth CI] 1.6, 31 times, z-value = 0.00796). Reaction times data also show responses on contrasts between a major place and an intermediate place take 86ms longer than responses on contrasts between two major places (95% CI 8, 164ms, t-value = 2.162).

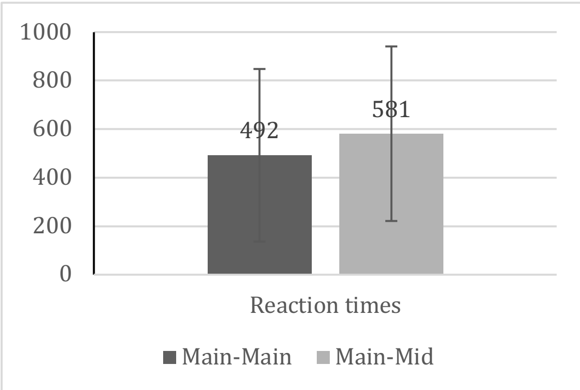

Figure 12. Mean reaction times and st. dev. for the contrasts between two main places (main-main) vs a main area and an intermediate area (main-mid) in the ABX task

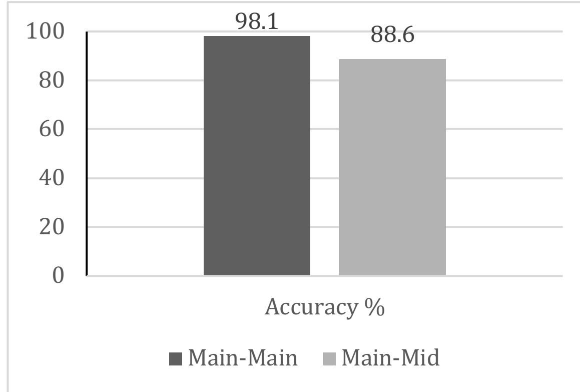

Figure 13. Mean accuracy rates for the contrasts between two main places (main-main) vs a main area and an intermediate area (main-mid) in the ABX task

Table 3 reports rates for 'same' responses and reaction times in ms for the same/different task, in which participants judged whether nonce signs in carrier phrases were the same or different.

Contrast	Same %	Rt(ms)	N of tokens
Temple - Forehead	12.5%	675	16
Forehead - 3	6%	705	15
Temple - 3	19%	534	16
Mouth corner - Cheek	0%	602	15
Mouth corner - 8	40%	723	15
Cheek - 8	25%	586	16
Mouth – Mouth corner	6%	521	16
Chin - Cheek	0%	446	15
Chin - 14	13%	635	15
Cheek -14	0%	554	16
Mouth - Chin	7%	478	14
Mouth - 13	40%	610	15
Chin - 13	93%	779	15
Nose - Mouth	0%	452	16
Nose - 11	6%	770	16
Mouth - 11	12.5%	635	16
Cheek - Temple	0%	482	15
Cheek - 6	46%	461	13
Temple - 6	12.5%	643	16
Ear - Temple	6%	560	16
Cheek - Nose	0%	442	16
Ear – Cheek	0%	699	14
Nose - Temple	0%	647	16
Cheek - Mouth	0%	631	16
Chin – Mouth corner	0%	465	15

Nose - Forehead	0%	609	14
Nose – Mouth corner	0%	586	14

Table 3. Rates of 'same' responses and reactions times (ms) for the same/different task

Rates of sameness are generally low, indicating that signers could discriminate all stimuli well. However, higher rates of 'same' responses are usually found in contrasts between a major area and an intermediate one. This is illustrated in Figures 14 and 25, which reports mean reaction times and 'sameness' rates per type of contrast (main-main vs main-intermediate), respectively. This is also confirmed by statistical analyses. Contrasts between a major and an intermediate area are more likely to be rated 'same' than contrasts between two major areas (95% CI 1, 1.18 times, z-value < 0.001). Regarding reaction times, contrasts between a major and an intermediate area take 88ms longer to identify than contrasts between two major areas (95% CI 24, 153 ms, t-value = 2.699).

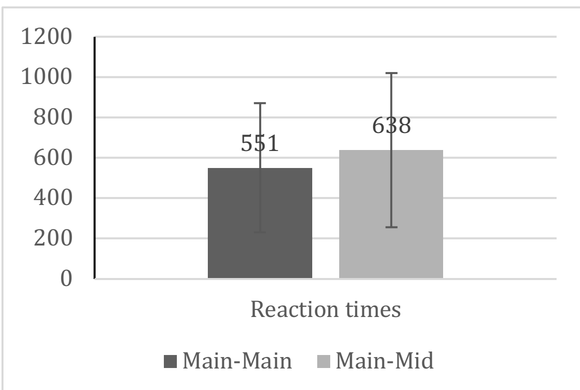

Figure 14. Mean reaction times and st. dev. for the contrasts between two main places (main-main) vs a main and an intermediate area (main-mid) in the same/different task

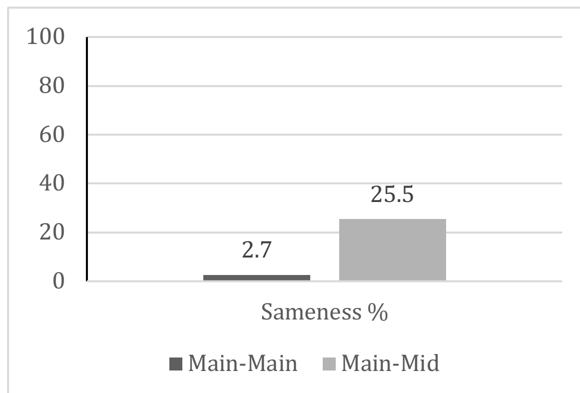

Figure 15. Mean 'sameness' rates for the contrasts between two main places (main-main) vs a main and an intermediate area (main-mid) in the same/different task

4.3. DISCUSSION. The perception experiment featured an ABX task with nonce signs, which targeted contrasts between neighboring locations and their intermediate points. It also presented participants with a same/different task, in which they judged whether two carrier phrases with nonce signs minimally different in terms of location were the same or different. Signers performed generally well in the two tasks, as they were able to discriminate between major areas, which were posited as potentially phonologically contrastive in the elicitation experiment, but also between a major area and an intermediate area. These results are not surprising, and overall agree with what previous literature on sign location discrimination found (Newport 1982, Emmorey et al. 2003, Morford et al. 2008).

Despite this, both tasks demonstrate that signers are more accurate and faster at discriminating between two major areas than between a major area and an intermediate area. According to Emmorey et al. (2003), peaks in accuracy for locations and handshapes are expected when stimuli straddle a category boundary. It follows that the difference in performances between main area contrasts and main-intermediate contrasts can only be explained if we assume that contrasts between major areas straddle phonological category boundaries, while contrasts between a major area and an intermediate area are within category boundaries. This is also supported by the elicitation data in the first experiment which used actual ASL signs, as all eight tested areas were phonetically distinct based on the statistical analyses of the obtained coordinate values.

The results in the perception experiment are consistent with the hypothesis that the main areas found in the elicitation experiment may be phonologically contrastive in ASL. These areas are: ‘forehead’, ‘temple’, ‘cheek’, ‘ear’, ‘chin’, ‘mouth’, ‘nose’, ‘mouth corner’. One limitation of the perception experiment is that, at present, the amount of data collected is not enough to draw reliable conclusions about individual contrasts. Nonetheless, the overall pattern is strong enough to be statistically significant, so we can tentatively assume that the majority of contrasts between the tested main areas are perceived by signers as phonemic. Moreover, qualitative data analysis reported next can provide more valuable suggestions as to which locations are likely to be phonemic in ASL.

4.3.1. QUALITATIVE ANALYSIS. The analysis presented below focuses on accuracy rates (from the ABX task) and rates of ‘sameness’ (from the same/different task).

The contrast between ‘mouth’ and ‘chin’ reflects exactly the results of the elicitation experiment. The contrast between the main areas has a 100% accuracy rate and very low ‘sameness’ rate, while the contrast between ‘chin’ or ‘mouth’ and their intermediate position (#13) has exceptionally low rates of accuracy and very high rates of ‘sameness’. The extreme rates of accuracy and ‘sameness’ ‘chin’ or ‘mouth’ and their intermediate position are explained by the large amount of overlap found in the elicitation experiment for the two areas, unlike what was found for other areas (see Figures 1 and 2). Despite the overlap, the elicitation experiment finds a statistical difference between the two areas, and the perception experiment has 100% accuracy rate and a very low ‘sameness’ rate, suggesting that these are distinct categories. This would go against what Stokoe et al. (1960), Stokoe (1965), and Sandler (1989) have proposed, as they conflated the two categories in their frameworks, but it would agree with Kegl & Wilbur (1976), Wilbur (1979), and Brentari (1998), as they distinguished ‘mouth’ and ‘chin’ as phonemic places of articulation.

The contrast between ‘chin’ and ‘cheek’ is well defined, with 100% accuracy rate in the ABX task and 0% sameness rate. Of course, all previous accounts have also stated that ‘chin’ and ‘cheek’ are two different places of articulation (Stokoe et al. 1960, Stokoe 1965, Kegl & Wilbur 1976, Friedman 1977, Wilbur 1979, Sandler 1989, Brentari 1998). The contrast with the intermediate position (#14) is of interest, because any error in discrimination or judgment as ‘same’ happened in the contrast ‘chin-14’ and not ‘cheek-14’, suggesting that the intermediate position is probably within the ‘chin’ categorical boundary but not the ‘cheek’ one.

‘Temple’ also seems to constitute its own phonological category. It is pretty distinct from ‘cheek’, with a 94% accuracy rate and 0% ‘sameness’ rating. Their intermediate point (#6) seems to share the same category boundary as ‘cheek’, as it has lower accuracy rating and much higher ‘sameness’ rating with ‘cheek’ than with ‘temple’. On the other hand, the contrast between ‘temple’ and ‘forehead’ is not as clear-cut. First of all, some signs, such as CLEVER, can be produced either on the forehead or the temple. Additionally, although the elicitation experiment found a significant difference between ‘forehead’ and ‘temple’, the overlap in the heatmaps is quite substantial (see Figures 4-5). The latter does not seem to be problematic for signers, as the contrast between the two areas has 100% accuracy rate and low rate of sameness.

Some signs, such as DREAM and CONCEPT may be produces either at the forehead or temple. This does not necessarily mean that the two areas are within the same phonological category. In fact, this case is not dissimilar to what happens in some spoken languages like Italian, in which the phonemes /e/ and /ɛ/ (and /o/ and /ɔ/) can be used virtually interchangeably in stressed syllables and vary regionally, even though they are distinct phonemes (Krämer 2009). Therefore, there is some merit, based on the experimental data solicited in the present study, in arguing for 'temple' being a phonological category. Such claim is in line only with Kegl & Wilbur (1976) and Wilbur (1979), who mention the category 'upper cheeks', which might be equated with 'temple'.

The area 'nose' can also be assumed to constitute a phonological category. Not only was it distinct in the elicitation experiment, it also had a 100% accuracy rate when contrasted with 'mouth' and 0% sameness rates when contrasted with all its neighboring areas: 'mouth', 'temple', cheek, 'forehead', and 'mouth corner'. The existence of the area 'nose' as a phonological category is mostly uncontroversial in previous literature, with the exception of Brentari (1998), who conflates the categories 'nose' and 'cheek'. However, there is nothing in the data here collected that suggests that 'nose' and 'cheek' are not two distinct categories.

Finally, the areas 'ear' and 'mouth corner' are also fairly distinct from other areas. The first, although not included in the ABX task, is significantly different from its neighboring areas in the elicitation experiment, and presents very low, 0% and 6%, 'sameness' rates in the contrasts with 'cheek' and 'temple' respectively. Likewise, 'mouth corner' is significantly different from neighboring areas in the elicitation experiment, has high rates of accuracy in the contrasts with 'cheek' and 'mouth, and very low rates of 'sameness' in the contrasts with 'mouth', 'chin' and 'cheek'. Proposing that these two areas are phonological categories might be considered controversial, as no authors have previously considered 'mouth corner' a phonological category, and only Kegl & Wilbur (1976) and Wilbur (1979) consider 'ear' distinctive. However, the perception experiment suggests that signers are very aware of them and have developed category boundaries.

5. Conclusions. The elicitation experiment found that the eight locations in ASL investigated in the present study, 'forehead', 'temple', 'cheek', 'ear', 'nose', 'mouth', 'chin', and 'mouth corner', are all produced in the areas of the face which are spatially different, as confirmed by the statistical analysis of the coordinate values. The perception experiment data support that these areas might be indeed phonological categories, on the basis of higher accuracy rates, faster responses, and lower rates of 'sameness' in the contrasts between two of these major areas than in the contrasts between a major area and an intermediate area.

Although additional research is needed to confirm the aforementioned assumptions, this study offers a first attempt at investigating facial contrasts other than APPLE-ONION and establishing with empirical research which facial locations may be phonemic in ASL.

Future studies should extend this line of inquiry by using more tokens per contrast, in order to have a more fine-grained statistical analysis shedding light on the nature of the individual contrasts tentatively proposed in the present study. Moreover, a replication of the elicited production experiment would ideally collect actual naturalistic production data, since the latter could be more ecologically valid, even if using the same face in this study provided the advantage of results more easily quantifiable. It should also be noted that this study did not investigate all possible phonemic locations on the head; the status of locations such as 'top of the head', 'eyes', 'side nose', 'under nose', and 'brows' remains to be addressed.

References

Baker, Stephanie, J. William Idsardi, M. Roberta Golinkoff & Laura-Ann Petitto. 2005. The perception of handshapes in American Sign Language. *Memory & Cognition* 33. 887–904. https://doi.org/10.3758/BF03193083.

Battison, Robbin. 1978. *Lexical borrowing in American Sign Language*. Silver Spring: Linstok Press.

Battison, Robbin, Harry Markowicz & James Woodward. 1975. A good rule of thumb: Variable phonology in American Sign Language. In Ralph W. Fasold & W. Roger Shuy (eds.), *Analyzing variation in language*, 291–302. Washington: Georgetown University Press.

Best, Catherine T., Gaurav Mathur, A. Karen Miranda & Diane Lillo-Martin. 2010. Effects of sign language experience on categorical perception of dynamic ASL pseudosigns. *Attention, Perception, & Psychophysics* 72(3). 747–762.

Boersma, Paul. 2012. Modelling phonological category learning. In Abigail C. Cohn, Cécile Fougeron & K. Marie Huffman (eds.), *Handbook of laboratory phonology*, 207–218. Oxford: Oxford University Press.

Brentari, Diane. 1998. *A prosodic model of sign language phonology language, speech, and communication*. Cambridge: MIT Press.

Brentari, Diane. 2002. Modality differences in sign language phonology and morphophonemics. In Richard P. Meier, Kearsy Cormier & David Quinto-Pozos (eds.), *Modality and structure in signed and spoken languages*, 35–64. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486777.

Brentari, Diane. & Petra Eccarius. 2011. When does a system become phonological?: Potential sources of handshape contrast in sign languages. In Rachel Channon & HarryVan der Hulst (eds.), *Formational units in sign languages*, 125–150. Berlin/Boston: De Gruyter. https://doi.org/10.1515/9781614510680.

Carney, Arlene E., P. Gregory Widin & F. Neal Viemeister. 1977. Noncategorical perception of stop consonants differing in VOT. *The Journal of the Acoustical Society of America* 62(4). 961–970. https://doi.org/10.1121/1.381590.

Corina, David. 1990. Reassessing the role of sonority in syllable structure: Evidence from a visual-gestural language. In Michael Ziolkowski, Lise Dobrin & K. Deaton (eds.). *Papers from the Chicago Linguistic Society*, 33–44. Chicago: Chicago Linguistic Society.

Eccarius, Petra. 2008. *A constraint-based account of handshape contrast in sign language*s. West Lafayette: Purdue University dissertation.

Friedman, Lynn. 1977. *On the other hand: New perspectives on American Sign Language*. Cambridge: Academic Press.

Fry, D. B., Arthur Abramson, Peter Eimas & Alvin Liberman. 1962. The Identification and Discrimination of Synthetic Vowels. *Language and Speech* 5(4). 171–189. https://doi.org/10.1177/002383096200500401.

Emmorey, Karen, Stephen McCullough & Diane Brentari. 2003. Categorical perception in American Sign Language. *Language & Cognitive Processes* 18. 21–45. https://doi.org/10.1080/01690960143000416.

van der Hulst, Harry & Els van der Kooij. 2006. Phonetic implementation and phonetic prespecification in sign language phonology. In Goldstein, Louis, D.H. Whalen & T. Catherine Best (eds.), *Papers in Laboratory Phonology 8*. 265-286. Berlin/New York: de Gruyter. https://doi.org/10.1515/9783110197211.

Kegl, Judy & Ronnie Wilbur. 1976. Where does structure stop and style begin? In Salikoko S. Mufwene, Carol A. Walker & Sanford B. Seever (eds.), *12th Regional Meeting of the Chicago Linguistic Society (CLS),* 376–396. Chicago: Chicago Linguistic Society.

Krämer, Martin. 2009. *The phonology of Italian*. Oxford: Oxford University Press.

Liberman, Alvin, Katherine Safford Harris, Howard Hoffman & Belver Griffith. 1957. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54(5). 358–368.

Marian, Viorica, K. Henrike Blumenfeld & Margarita Kaushanskaya. 2007. The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research* 50. 940–967. https://doi.org/10.1044/1092-4388(2007/067).

Martinez, Marc G. 2015. *Categorical perception in Catalan Sign Language*. Barcelona: Universitat Pompeu Fabra MA thesis.

Morford, Jill, Angus Grieve-Smith, James MacFarlane, Joshua Staley & Gabriel Waters. 2008. Effects of language experience on the perception of American Sign Language. *Cognition* 109. 41-53.

Newell, William, Frank Caccamise, Keitha Boardman & Barbara Ray Holcomb. 1983. Adaptation of the Sign Language Proficiency Interview (SLPI) to assessing sign communicative competence. *Sign Language Studies* 72(41). 311–353.

Newport, Elissa. 1982. Task specificity in language learning? Evidence from speech perception and American Sign Language. In Eric Wanner & R. Lila Gleitman (eds.), *Language acquisition: The state of the art*, 450–486. Cambridge: Cambridge University Press.

Peirce, Jonathan. 2007. PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods* 162. 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017.

Pisoni, David. 1975. Auditory short-term memory and vowel perception. *Memory & Cognition* 3(1). 7–18. https://doi.org/10.3758/BF03198202.

Pisoni, David. & Jeffrey Tash. 1974. Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics* 15(2). 285–290. https://doi.org/10.3758/BF03213946.

Popper, Roger. 1972. Pair discrimination for a continuum of synthetic voiced stops with and without first and third formants. *Journal of Psycholinguistic Research* 1(3). 205–219.

Qualtrics. 2005. Qualtrics [computer software]. https://www.qualtrics.com.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/.

Sandler, Wendy. 1989. *Phonological representation of the sign: Linearity and nonlinearity in sign language phonology*. Dordrecht: Foris.

Sandler, Wendy. 2012. The phonological organization of sign languages. *Language & Linguist Compass* 6(3). 162–182. https://doi.org/10.1002/lnc3.326.

Stauffer, Linda. 2011. ASL students' ability to Self Assess ASL Competency. *Journal of Interpretation* 21(1). 80-95.

Sternberg, Martin. 1998. *American Sign Language Dictionary* (3rd Ed.). New York: Harper Perennial.

Stokoe, William. 2005 [1960]. Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf. *The Journal of Deaf Studies and Deaf Education* 10(01). 3–37.

Stokoe, William, Dorothy Casterline & Carl Croneberg. 1965. *A dictionary of American sign language on linguistic principles*. Silver Spring: Linstok Press.

Tartter, Vivienne. 1981. A comparison of the identification and discrimination of synthetic vowel and stop consonant stimuli with various acoustic properties. *Journal of Phonetics* 9(4). 477–486. https://doi.org/10.1016/S0095-4470(19)31022-8.

Wilbur, Ronnie. 1979. *American Sign Language and sign systems*. Baltimore: University Park Press.