2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Evaluation of the Potential of Automatic Naming Latency Detection for Different Initial Phonemes during Picture Naming Task*

Sunghea Park, Sven Altermatt, Sandra Widmer Beierlein, Anja Blechschmidt, Claire Reymond, Markus Degen, Eliane Rickert, Sandra Wyss, Katrin P. Kuntner, and Simone Hemm, *Member, IEEE*

***Abstract*— Naming latency (NL) represents the speech onset time after the presentation of an image. We recently developed an extended threshold-based algorithm for automatic NL (aNL) detection considering the envelope of the speech wave. The present study aims at exploring the influence of different manners (e.g., "m" and "p") and positions (e.g., "t" and "p") of articulation on the differences between manual NL (mNL) and aNL detection.**

Speech samples were collected from 123 healthy participants. They named 118 pictures in German, including different initial phonemes. NLs were manually (Praat, waveform and spectrogram) and automatically (developed algorithm) determined. To investigate the accuracy of automatic detections, correlations between mNLs and aNLs were analyzed for different initial phonemes.

ANLs and mNLs showed a strong positive correlation and similar tendencies in initial phoneme groups. ANL mean values were shorter than the ones of mNLs. Nasal sounds (e.g., /m/) showed the largest and those for fricatives (e.g., /s/) the smallest difference. However, in fricatives, 39% of NLs were detected later by automatic detections than by manual detections, which led to a reduced mean difference with mNLs. The signal energy of the initial phonemes, i.e., if they are voiced or voiceless, influences the form of the speech envelope: initial high signal energy is often responsible for an early detection by the algorithm.

Our study provides evidence of a similar tendency in mNL and aNL according to different positions of articulation in each initial phoneme group. ANLs are highly sensitive to detection of speech onsets across different initial phonemes. The dependency of the NL differences on the initial phonemes will lose importance during progress evaluations in aphasia patients if the relative changes for each picture are considered separately. Nevertheless, the algorithm will be further optimized by adapting its parameters for each initial phoneme group individually.

***Clinical Relevance*— This underlines the feasibility to use automatic naming latency detection for the evaluation of patients with aphasia in a clinical setting as well as for practices at home during picture naming.**

I. Introduction

Naming latency represents the speech onset time predicting a temporal process of word and phonological retrievals as well as speech production [1, 2]. Particularly, speakers with aphasia showed longer NLs than healthy speakers. Aphasia is referred to as language disorder which can affect understanding, reading, writing and speaking [3]. Word finding problems are the core symptom of all types of aphasia and have therefore a big influence on naming latencies. Our research group recently developed an extended threshold-based algorithm for aNL detection considering the envelope of the speech to be integrated in a mobile application for clinical use. The influence of different manners and positions of articulation on the differences between mNL and our aNL detection has not yet been investigated.

NLs have been investigated as a parameter to predict a process of word retrieval, phonological encoding and speech production by measuring a speech onset time following a naming task [1, 4, 5]. The detection of naming latency can be made manually or automatically. Determining the NL manually is currently the gold standard and is set as the target for aNL detection algorithms. In previous studies, delays and errors of automatic detections of voice keys have been reported in measuring acoustic naming latencies [2, 6–8]. However, the development of automated software algorithms has improved sensitivity and accuracy of speech onset detections, which aims to replace efforts of manual work in a large data set [9, 10]. The open-source tool "Chronset", published by Roux, Armstrong and Carreiras [10], uses in addition to the time signal also a time-frequency spectrogram. Out of these two signals, the time signal and the spectrogram,

*Research supported by University of Applied Sciences and Arts Northwestern Switzerland.

S. Park is with the Institute for Special Education and Psychology, School of Education, University of Applied Sciences and Arts Northwestern Switzerland, Switzerland (sunghea.park@fhnw.ch).

S. Altermatt is with the Institute for Medical Engineering and Medical Informatics, School of Life Sciences, University of Applied Sciences and Arts Northwestern Switzerland, Switzerland (sven.altermatt@fhnw.ch)

S. Widmer Beierlein is with the Institute for Special Education and Psychology, School of Education, University of Applied Sciences and Arts Northwestern Switzerland, Switzerland (sandra.widmer@fhnw.ch).

A. Blechschmidt is with the Institute for Special Education and Psychology, School of Education, University of Applied Sciences and Arts Northwestern Switzerland, Switzerland (anja.blechschmidt@fhnw.ch).

C. Reymond is with the Institute for Visual Communication, Academy of Art and Design, University of Applied Sciences and Arts Northwestern Switzerland, Switzerland (claire.reymond@fhnw.ch).

M. Degen is with the Institute for Medical Engineering and Medical Informatics, School of Life Sciences, University of Applied Sciences and Arts Northwestern Switzerland, Switzerland (sven.altermatt@fhnw.ch)

E. Rickert is with the Institute for Medical Engineering and Medical Informatics, School of Life Sciences, University of Applied Sciences and Arts Northwestern Switzerland, Switzerland (eliane.rickert@fhnw.ch)

S. Wyss is with the Institute for Medical Engineering and Medical Informatics, School of Life Sciences, University of Applied Sciences and Arts Northwestern Switzerland, Switzerland (sandra.wyss@fhnw.ch)

K. Kuntner is with the Institute for Special Education and Psychology, School of Education, University of Applied Sciences and Arts Northwestern Switzerland, Switzerland (katrinpetra.kuntner@fhnw.ch).

S. Hemm is with the Institute for Medical Engineering and Medical Informatics, School of Life Sciences, University of Applied Sciences and Arts Northwestern Switzerland, Switzerland (corresponding author, phone: +41 61 228 56 89; fax: +41 61 467 47 01; email: simone.hemm@fhnw.ch).

six different features are generated. Every feature has its own time course and threshold. As soon as four of the six features of a specific signal exceed the threshold for more than 35ms, the NL is detected. The algorithm developed by Jansen & Watter [9] frames the time signal and contains five consecutive steps where the signal is analyzed in different ways. The main goal of all these steps is to differentiate parts of the signal with noise and parts with human speech with the help of different heuristic approaches: The signal is analyzed in different ways and if one step fails to detect speech or noise there are four more safety nets. The SayWhen algorithm also flags files which need to be manually reviewed because of uncertainty. To the best of our knowledge none of these algorithms has been integrated so far into a mobile application.

Some researchers mentioned that different initial phonemes can influence automatic detections of naming latency when an automatic detection of voice key is used [11, 12]. Particularly, the weak energy of initial phonemes such as voiceless fricatives, e.g., /f, s/, can be missed or be detected late in the automatic detection [2, 6]. Voiceless consonants, particularly the fricatives, produce a weak high frequency energy because they do not contain a low and clear frequency energy of voicing. As initial phonemes contain various phonetic and acoustic characteristics, each initial phoneme can represent different speech onset times according to the manner and position of articulation as well as to the vibration of vocal folds [2, 13]. Sakura and colleagues reported that aNL detections were more than 100 milliseconds later than mNLs. According to a study of reading monosyllables [14], mNLs were influenced by the manner of articulation (see Tab.1 for an overview of different manners), in which plosives (e.g., /p/, /t/, /g/) showed the longest latency than fricatives (e.g., /s/, /f/) and nasals (e.g., /m/, /n/). Voiced plosives (e.g., /g/) and fricatives had longer NLs than voiceless consonants (e.g., /p/). In addition, they presented the influence of the articulatory position in naming latency, in which alveolar positions were detected earlier than labiodental and interdental positions. From a clinical perspective, NLs of persons with aphasia (PWA) were reported to be longer than the one of people without aphasia. Long naming latency indicates a difficulty of semantic word retrieval as well as a difficulty of phonological encoding [15].

Although automated algorithms have been enhanced in detecting naming latency, it has not been yet investigated how different characteristics of initial phonemes influence slower or earlier automatic detections compared with manual detections. The study aims to explore the effect of different manners and positions of initial phonemes on our automated detection algorithm. In order to do so, aNLs have been compared with mNLs, the gold standard, by categorizing initial phoneme's subgroups.

II. Material and Method

A. Subjects

123 speakers (50 males, 72 females and 1 unknown) with the average age of 42.28 (range between 18 and 82 years) participated in the naming experiment. All speakers were bivarietal and spoke Swiss German (dialect) as first language and have learned Standard German (standard variety) at school. All speakers did not have speech and hearing disorders. All participants gave informed consent for the use of their audio and video data for research.

B. Data collection

Data collection was performed within a study testing images for name agreement using a specifically developed application. The naming task consisted in naming images with a single word beginning with a consonant. 118 images of single words (62 nouns and 56 verbs) with different subgroups of initial phonemes (see Table 1) were selected for the study. Each image was presented one by one to the participants in the following way: first a fixation cross was shown for 500ms to direct the attention of the participant to the screen. Then, a black screen appeared for 150ms before the image was shown. After the naming, the speaker could move to the next image, starting again with the fixation cross. In order to avoid a bias of item's order in data collection, 8 different sets of image order were randomly applied to participants. Videos and audios were recorded, and time stamps saved for an optimal synchronization between recordings and image presentation.

TABLE I. Manners of Articulation, the Initial Phonemes Being Part of Each Articulation Group and the Number of Recordings Available for Each Phoneme (Sub)Group.

Manner of Articulation	Position of Articulation			
Plosive (n=3262)	**bilabial** (n=1407)	**alveolar** (n=1018)	**velar** (n=837)	
	/p/, /b/	/t/, /d/	/k/, /g/	
Fricative (n=4482)	**labio-dental** (n=1054)	**alveolar** (n=1074)	**post-alveolar** (n=2127)	**glottal** (n=227)
	/f/, /v(w)/	/s/	/sch/	/h/
Affricate (n=661)	**labio-dental** (n=257)	**alveolar** (n=347)	**palato-alveolar** (n=57)	
	/pf/	/ts(z)/	/tsch/	
Approximant (n=1374)	**(Glide) palatal** (n=236)	**(Liquid) lateral** (n=546)	**(Vibrate) alveolar/uvular** (n=592)	
	/j/	/l/	/r/	
Nasal (n=580)	**bilabial** (n=376)		**alveolar** (n=204)	
	/m/		/n/	

Manner of articulation: Plosive: airflow is totally blocked, the air accumulates in the vocal tract and gets released in the form of a burst; Fricative: sound made by air streams through narrow channel which generates a constriction of oral cavity; Affricate: is the combination of a plosive and fricative, first the airflow is fully stopped and then released as a fricative Approximant: airflow escapes mouth with less disturbance compared to other manner, Liquid made by airflows with the sides of the tongue while Glide made like a vowel movement from one to the other place of articulation; Nasal: like plosive manner, airflow is completely blocked during releasing continuous airflow through the nose; *Position of articulation:* bilabial: formed by closure or near closure of the lips; alveolar: articulated with the tongue against or close to the superior alveolar ridge; velar: articulated with the back part of the tongue against the soft palate; labio-dental: articulated with the lower lip touching the upper front teeth; post-alveolar: as alveolar consonants, but farther back in the mouth; glottal: sound made at the glottis between vocal folds; palato-alveolar: articulated with the blade or tip of the tongue raising toward just behind of the alveolar ridge; palatal: produced by holding the tongue high in the mouth toward soft palate; lateral: partial closure in the middle of mouth by the tongue with airstream along the side of the tongue; uvular: articulated with the back of the tongue against or near the uvula, farther back in the mouth than velar consonants.

C. Data preparation

Speech samples of the study consisted of 10359 audio files, which included only correct responses of target words. Each image had a single target word, which was counted as a correct response in the study. Non target reactions as well as recordings with loud nonverbal noises, which disturbed the automated algorithm of speech detection, were excluded from the analysis.

D. Latency detection

NLs were manually and automatically detected. For the **manual detection**, each naming latency was calculated as an interval by measuring the onset of the correct target speech response following the onset of the picture naming stimulus on the tablet screen. Naming latencies were measured using a speech analysis software, Praat[1], which enables acoustic speech analysis by presenting wave form and spectrogram of audio data.

For the **automatic detection** of the naming latency, an extended threshold-based approach, was implemented in Matlab[2] considering three parameters. The first one is the threshold value, above which the signal is considered for analysis. The second one describes the duration of the envelope of the speech wave for which it must be over the threshold until it is considered as a full word. This allows to exclude short noises like sneezing or coughing. During human speech, the envelope of the speech wave can fall under the threshold for a certain time especially in long words. In consequence, a third parameter was implemented characterizing the time, the signal is allowed to be under the threshold before the end of the word. These three parameters were optimized with the available speech data from the study using the Nelder-Mead algorithm (*fminsearch* function from Matlab). The resulting values were 111.8ms for the minimum time the envelope must be over the threshold, 457.4ms for the maximal time under the threshold and 15.9% of the amplitude for the threshold itself.

E. Statistical Analysis

To investigate the accuracy of automatic detections, absolute mNLs and aNLs were statistically compared for the different initial phoneme groups and subgroups by analyzing correlations between mNL and aNL. Furthermore, the differences between mNL and aNL were calculated and statistically compared for the different subgroups of initial phonemes as indicated in Table 1. To identify significant differences, Wilcoxon signed rank test was used. To measure the strength and directions of association between aNL and mNL, Kendall's tau-b correlation analysis was applied by using IBM SPSS Statistics (version 26). In addition, examples of each phoneme group were chosen, and the speech signal visualized together with the identified threshold, the detected speech envelope as well as with the mNL and aNL. The difference between the two latencies were visually analyzed considering the initial phoneme.

III. Results

NLs detected automatically (1151.59ms±639.64ms) and manually (1193.62ms±643.15ms) showed a strong positive correlation (Kendall's tau-b, TB=0.88, p<0.0005). The different subgroups clearly show similar tendencies by subgroups (Fig. 1). While NLs of fricatives are in a similar range, plosives, affricates, approximants and nasals show higher variations within the subgroups. But the position of the articulation seems to make a difference as bilabial sounds have longer NLs than alveolar or velar sounds in each initial phoneme group.

For the initial phoneme subgroups, the correlations of mNL and aNL were strong: plosives (TB=0.899), fricatives (TB=0.867), affricates (TB=0.912), approximants (TB=0.916) and nasals (TB=0.897) all at significant levels (p<0.0005).

Distributions of negative and positive differences between mNLs and aNLs are shown in Fig. 2. ANLs were significantly different from mNLs (z=-57.694, p<0.0005): plosives (z=-44.919), fricatives (z=-16.5), affricates (z=-12.294), approximants (z=-30.14) and nasals (z=-20.821) all at significant levels (p<0.0005). While nasal sounds showed the largest mean difference (101.68ms±55.95ms) and those for fricatives the smallest one (16.66ms±68.94ms), the algebraic sign of the difference has to be considered to evaluate the

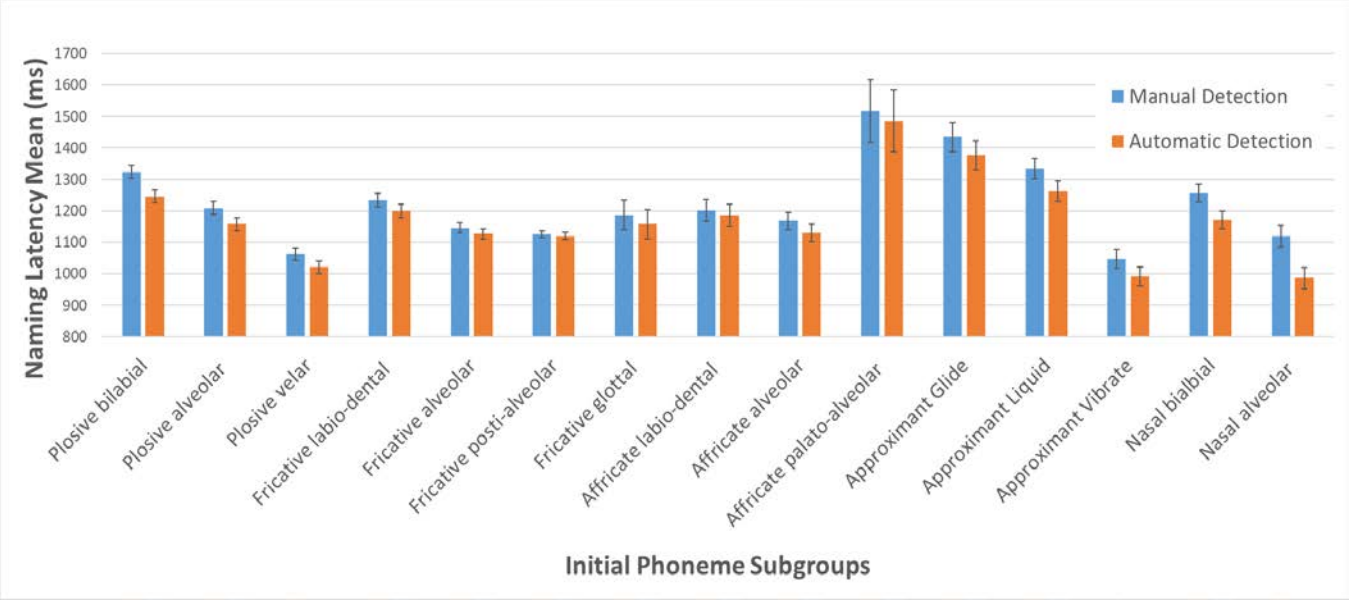

Figure 1: Mean of aNL and mNL in the different initial phoneme groups (error bars: +/- 2 standard error of the mean)

[1] https://www.fon.hum.uva.nl/praat/, [2] https://mathworks.com

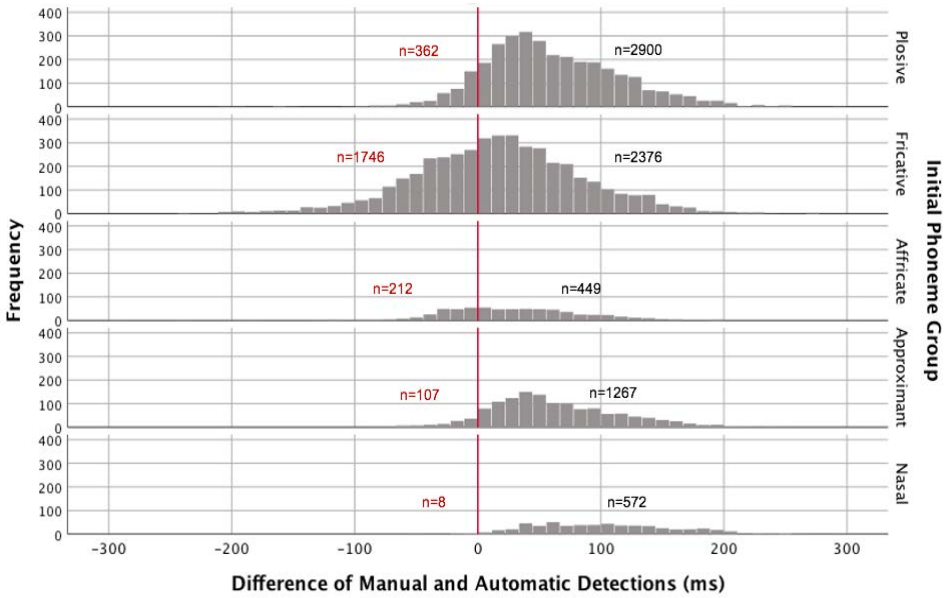

Figure 2: Distribution of the differences between aNLs and mNLs. Negative values indicate that the mNL was smaller than the aNL, i.e., that the algorithm detected the beginning of the word later than the speech therapist.

impact on the mean values (Fig. 2). About 23.5% of the data showed negative differences, mostly in fricatives (72%) (see Fig. 2). In detail, voiceless fricatives (e.g., /s/) accounted for 97% of the negative difference in the fricative and voiceless plosives in contrast only for 75%. Such negative values led to reduced mean differences between mNLs and aNLs. When considering the differences between the manual and the automatic detection, one can see that in all groups the algorithm detected the beginning of the word earlier than the speech therapist. Only for fricatives the speech therapist often attributed the beginning of the word to an earlier moment than the algorithm.

Some typical and some extreme examples for the five initial phoneme groups are given in Fig. 3. Figure parts in the same row correspond to the same group: Plosives (Fig. 3 A. and B.), Fricatives (Fig. 3 C. and D.), Affricates (Fig. 3 E. and F.), Approximants (Fig. 3 G. and H.) and Nasals (Fig. 3 I. and J.). For each group one negative and one positive difference between the NLs is presented (Fig 3. A.-H.) except for the nasals for which nearly only positive differences exist (I., J.). Each graph shows the speech signal, the identified speech envelope, the threshold as well as mNLs and aNLs. Differences between mNL and aNL are indicated in each subfigure.

In general, the envelope fits better for voiceless initial phonemes (Fig. 3 A., C.), double consonants (Fig. E) as well as for backward position sound (Fig. G.), i.e., with less signal energy than for voiced phonemes (Fig. 3 B., D.), single consonant (Fig. F.) and the frontal position (Fig. H.). While the signal to noise ratio did not have an influence on the aNL detection, outliers can have an influence on the calculated envelope (Fig. 3 H.). The beginning of the envelope is not only influenced by the initial but also by the following phonemes (Fig. 3 D.). For the nasals, which include only voiced phonemes, the automatic detection is nearly always earlier than the manual one. The bilabial nasal sound is likely to be described later than the alveolar sound in both detections (Fig.3 I., J.).

IV. Discussion

Our results show that aNLs are sensitive to detection of speech onsets for different initial phonemes, but the tendency of manual and automatic detection is similar.

Previous studies [12, 15] reported about the influence of manners and positions of initial phonemes on NLs as well as on the impact of the characteristics of different phonemes on both automatic and manual detections. Automatic detections, in previous studies, showed different tendencies from manual detections depending on phonemes due to automatic detection error. The position of articulation is another aspect to make a difference of NLs. For instance, bilabial sounds showed the longest NLs, following alveolar and velar initial phonemes of plosives in both detections. Alveolar positions were detected earlier than labio-dentals in fricatives and affricates. Further, voicing is an element to influence NLs, in which voiceless consonants had longer NLs than voiceless sounds in our study as well. In contrast to the previous studies, our automatic detection showed similar tendencies with the manual detection according to the different positions of initial phonemes. In addition, the automatic detection was, in average, 42ms earlier than the manual detection in considering only mean values. However, as presented in Figure 2, 23.5% of our data were detected later in the automatic algorithm, mostly in fricatives, while 76.5% was descried earlier in the algorithm than in the manual. Likewise, NLs of our data presented longer than those of previous studies with reading monosyllables because of the picture naming task with polysyllabic words.

Another interesting result of the study is a tendency of negative and positive differences of manual and automatic detections between voiced- /b, d, g, v(w)/ and voiceless consonants /p, t, k, f, h, s, sch/. In plosives, 75% of negative differences were voiceless consonants while 97% of negative differences was voiceless sounds in fricatives. The tendency indicates that the automatic algorithm detected, on one hand, voiceless initial phonemes slower than the speech therapist.

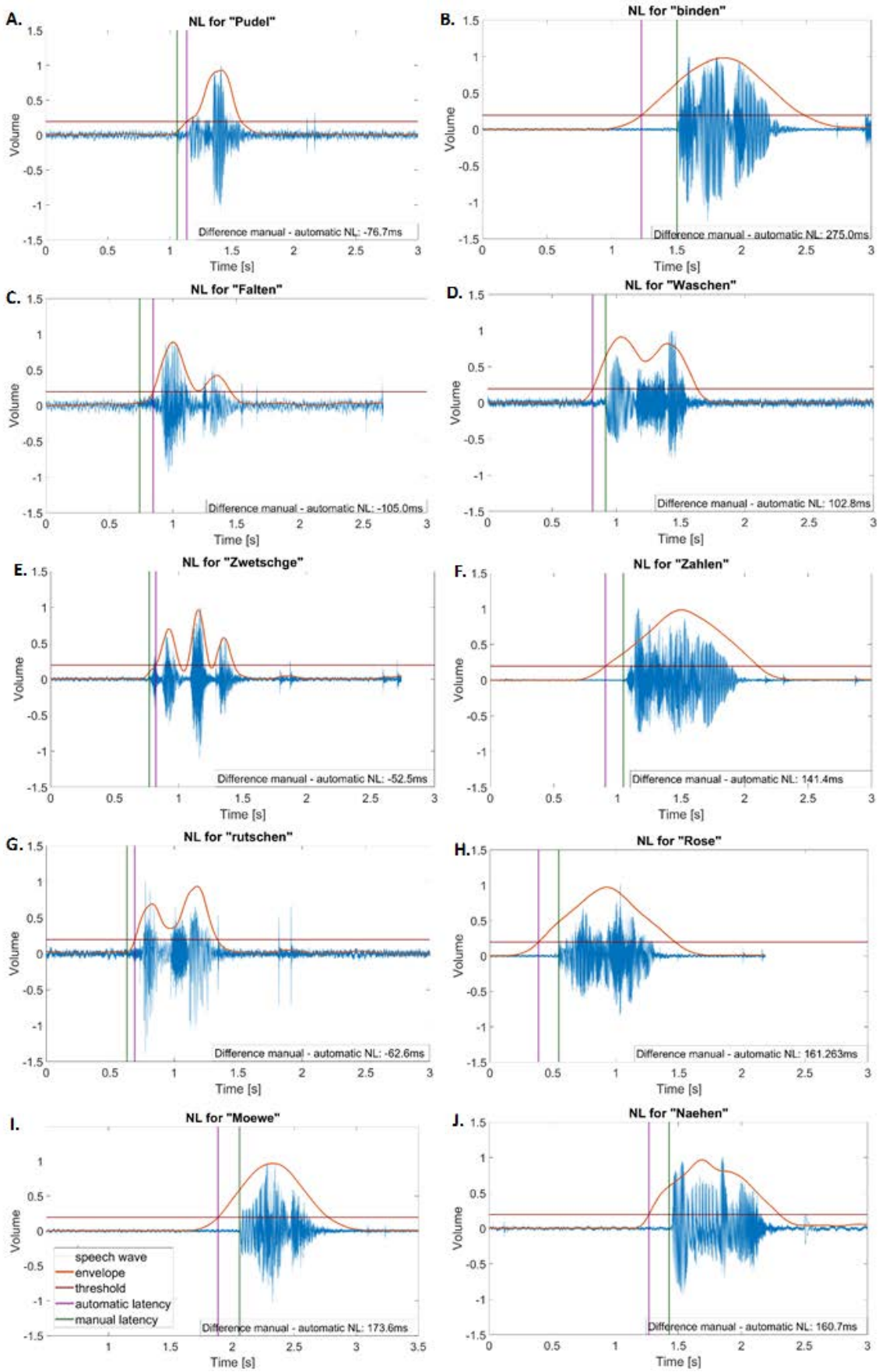

Figure 3: Examples for different phoneme groups: Plosive alveolar (A.; B.), Fricative labio-dental (C.; D.), Affricate alveolar (E.; F.), Approximant (uvular G.; alveolar H.), Nasal (bilabial I.; alveolar J.).

On the other hand, the algorithm descried voiced initial phonemes faster than the manual detection. The reason could be that the signals of voiceless consonants contain less or weaker intensity of energy in comparing with voiced sounds, which could impact on the algorithm to generate optimal envelopes differently. Additionally, double voiceless consonants /sp, st/ could make stronger and longer aspiration than single voiceless consonants in the beginning of a word, which could result in large negative differences of fricatives. The dependency of the NL differences on the initial phonemes will lose importance during a progress evaluation in aphasia patients considering NLs for each naming task separately.

ANL detection has already been investigated with different approaches. The already mentioned open-source tool "Chronset" [10] is available online or as Matlab source code. To evaluate their tool, a comparison between manual and automatic detected NL was performed (R^2 = 0.97, offset = 21ms; proportion of regression residuals within ±10ms range = 26%, SD = 90ms). Compared to the approach presented in this paper it is a very time-consuming procedure. The aNL detection algorithm SayWhen from Jansen and Watter [9] was originally implemented in Matlab and later transferred to Visual C++. The performance of the algorithm was tested compared to mNLs with 3940 files. 69.5% of the aNL's were within 10ms compared to the NLs. Although nearly half of the data were flagged which means 1838 results would have to be reviewed. This is a drawback with respect to the algorithm presented in this paper. Neither "Chronset" nor the SayWhen algorithm is adapted for the integration in a mobile application with aNL detection for therapeutical use while the presented algorithm is intended for such a use.

During further work, the algorithm will be improved in different areas to reduce the difference to the mNLs. One approach will be to the envelope function with the help of additional features considering the amount of noise such as the harmonic-to-noise ratio (HNR). With the help of this feature the filtering might be optimized. The lower the amount of filtering, the more similar is the slope of the envelope to the one from the original signal. Thus, too early detection can be reduced. A second promising approach is to adapt the values of the different parameters and to generate unique ones for every initial phoneme. Due to the fact, that the initial phoneme and the articulation manner and position for every expected word is known, this could help to improve the performance of the automatic detection algorithm. In addition to the improvements already mentioned, also the use of an automatic speech recognition (ASR) software will be investigated. The advantage would be to narrow the analyzed signal part down as the target word is recognized within the recording. An ASR algorithm will bring up the time point of the detection of onset of the word, thus further helps to automatically locate the NL.

V. Conclusion

Our study provides evidence of a similar tendency in manual and automatic detection of NLs with our extended threshold-based approach. Compared to already established aNL detection algorithms, the presented algorithm seems to be more robust regarding the standard deviation of the differences between mNL and aNL detection although the results showed an effect of the initial phonemes and the corresponding energies. The observed dependency of the NL differences on the initial phonemes will lose importance during a patient-specific progress evaluation in aphasia patients considering NLs for each named image separately. The next step will be to further improve the algorithm and to test it - implemented on a tablet prototype application – for the use of PWAs during picture naming exercises in a clinical setting as well as practices at home.

References

[1] P. Bonin, M. Chalard, A. Méot and M. Fayol, "The determinants of spoken and written picture naming latencies", *British Journal of Psychology*, vol. 93, pp. 89–114, 2002.

[2] B. Kessler and R. Treiman, "Phonetic Biases in Voice Key Response Time Measurements", *Journal of Memory and Language,* vol. 47, pp. 145-171, 2002.

[3] W. Huber, K. Poeck and D. Weniger, "Klinisch-neuropsychologische Syndrome und Störungen Aphasie", in *Klinische Neuropsychologie.* 6th ed. Thieme, 2006, pp. 93–173.

[4] M. Calabria, N. Grunden, M. Serra, C. Garcia-Sánchez and A. Costa, "Semantic Processing in Bilingual Aphasia: Evidence of Language Dependency", *Frontiers in Human Neuroscience*, vol. 13, pp. 1-15, 2019.

[5] E. E. Galletta and M. Goral, "Response Time Inconsistencies in Object and Action Naming in Anomic Aphasia", *American Journal of Speech-Language Pathology*, vol. 27, pp. 477–484, 2018.

[6] E. Bates, S. D'Amico, T. Jacobsen, A. Sze´kely, E. Andonova and A. Devescovi, "Timed picture naming in seven languages", *Psychonomic Bulletin & Review*, vol. 10, pp. 344–380, 2003.

[7] B. B. Holbrook, A. H. Kawamoto and Q. Liu, "Task Demands and Segment Priming Effects in the Naming Task", *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 45, no. 5, pp. 807-821, 2019.

[8] S. Mätzig, J. Druksa, J. Mastersonb and G. Viglioccoa, "Noun and verb differences in picture naming: Past studies and new evidence", *Cortex*, vol. 45, pp. 738-758, 2009.

[9] P. Jansen and S. Watter, "SayWhen: An automated method for high-accuracy speech onset detection", *Behavior Research Methods*, vol. 40, no. 3, pp. 744-751, 2008.

[10] F. Roux, B. C. Armstrong and M. Carreiras, "Chronset: An automated tool for detecting speech onset", *Behavior Research Methods*, vol. 49, pp. 1864-1881, 2017.

[11] K. Rastle and M. H. Davis, "On the complexities of Measuring Naming", *Journal of Experimental Psychology: Human Perception and Performance*, vol. 28, no. 2, pp. 307-314, 2002.

[12] A. Szekely, S. D'Amico, A. Devescovi, K. Federmeier, D. Herron, G. Iyer, T. Jacobsen and E. Bates, "Timed Action and Object Naming", *Cortex*, vol. 41, no. 1, pp. 7-25, 2002.

[13] N. Sakura, T. Fushimi and I. Tatsumi, "Measurement of naming latency of kana characters and words based on the speech wave analysis: Manner of articulation of a word-initial phoneme considerably affects naming latency", *Japanese Journal of Neuropsychology*, vol. 13, pp. 126–136, 1997.

[14] K. Rastle, K. P. Croot K, J. M. Harrington and M. Coltheart, "Characterizing the Motor Execution Stage of Speech Production: Consonantal Effects on Delayed Naming Latency and Onset Duration", *Journal of Experimental Psychology*, vol. 31, no. 5, pp. 1083-1095, 2005.

[15] J. Schuchard, E. L. Middleton and M. F. Schwartz, "The Timing of Spontaneous Detection and Repair of Naming Errors in Aphasia", *Cortex*, vol. 93, pp. 79–91, 2017.