Hands-free Devices for Displaying Speech and Language in the Tactile Modality - Methods and Approaches

Citation for published version (APA):

Kappers, A. M. L., & Plaisier, M. A. (2021). Hands-free Devices for Displaying Speech and Language in the Tactile Modality - Methods and Approaches. *IEEE Transactions on Haptics*, 14(3), 465-478. [9325066]. <https://doi.org/10.1109/TOH.2021.3051737>

DOI:

[10.1109/TOH.2021.3051737](https://doi.org/10.1109/TOH.2021.3051737)

Document status and date:

Published: 01/07/2021

Document Version:

Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Hands-free Devices for Displaying Speech and Language in the Tactile Modality - Methods and Approaches

Astrid M.L. Kappers and Myrthe A. Plaisier

Abstract—Communication is an essential part of human life. In this survey we give an overview of hands-free tactual devices that have been developed and tested for conveying speech or language. We opted for “hands-free” because especially in the case of individuals with impaired vision, in many situations their hands will be occupied with other essential tasks. We start this survey with presenting the various word building blocks that have been tested. These blocks vary from units based on the actual speech signal, via patterns representing phonemes, to letters, or letters coded via Morse or Braille-like patterns. In the second part of this survey, studies that use these building blocks to create words are discussed. General findings are that successful devices do not necessarily depend on underlying speech characteristics, dynamic patterns give better results than static patterns, and more vibrators do not generally give better results. Moreover, some of the most successful devices required only limited training time. Most of the recent devices are still in a quite early state of development and are tested only with a limited number of patterns. However, many of these recent devices give promising results and are worthwhile to further investigate and develop.

Index Terms—Language, speech, phonemes, letters, words, Morse

1 INTRODUCTION

THERE have been many attempts at conveying speech and/or language by means of tactual stimulation with a device. The underlying motivation of many of these studies is to develop a way of communication for and with people who cannot see and/or hear. As fingers are the most sensitive body parts for (vibro)tactile stimulation (e.g., [1], [2], [3], [4]), many previous studies designed and tested devices meant for use with the hands or fingers. Prominent successful examples are various types of Braille displays and at a more experimental stage are various kinds of Lorm gloves. However, individuals with blindness or deafblindness will often need their hands for other tasks. In this extensive literature overview we will focus on solutions that are suitable for those situations and therefore, we will only present hands-free devices for tactual communication.

For the design of tactual communication devices, it is very important to take human sensitivity into account. Vibrotactile sensitivity, for example, varies widely over the body [1], [4], and the discrimination threshold of vibrotactile frequencies depends on the frequency (e.g., [5], [6]). Importantly, performance also depends on type of vibrator used [7]. However, fundamental psychophysical performance lies outside the scope of the current overview, and the interested reader is referred to the literature (e.g., [4], [6], [8], [9], [10], [11], [12]).

The studies presented in this overview are all published

in peer-reviewed scientific journals or as a paper in conference proceedings. All references of the included papers have been checked for further relevant studies. In addition, papers that cite the included studies have also been checked for relevance. We only included papers that report on user studies, so that excludes papers that only present or analyse a device. Some papers report on several experiments, but we present here only the most relevant ones for the present purpose. Moreover, there are studies that investigated the value of a tactile device in combination with lipreading or auditory information. In such cases, we only report the baseline condition with the use of the tactile device alone. Devices specifically designed for navigation purposes are also not included as these focus on specific commands like “go to the right”, “go straightforward”, etc. It should be noted, however, that there exists an extensive literature on haptic navigation devices using belts or vests.

The studies presented in this overview vary widely in scope, duration (less than an hour – 100 hours), number of participants (2 – 45), and success. The scope includes, among others, devices that decompose the speech signal into different frequency bands that are subsequently used to activate a series of vibrators, devices that define different vibration patterns for different phonemes, devices that generate vibrotactile Morse code, and devices that simulate handwriting of letters. Not surprisingly, studies with large numbers of participants have often a much shorter duration than the ones that are actually case studies with just a few participants. Success rates also differ, but these are not always easy and fair to compare. Some studies investigate word or phoneme recognition with open response possibilities, whereas other studies test recognition rates on a (very) limited set of learned patterns or letters with a multiple choice paradigm. In most of the studies, hearing and sighted

- A. M. L. Kappers is with the Groups Dynamics and Control, Control Systems Technology, and Human Technology Interaction, Eindhoven University of Technology, 5612 AZ Eindhoven, The Netherlands (e-mail: a.m.l.kappers@tue.nl). Orcid: 0000-0003-4101-7717. M. A. Plaisier is with the Group Dynamics and Control, Eindhoven University of Technology, 5612 AZ Eindhoven, The Netherlands (e-mail: m.a.plaisier@tue.nl). Orcid: 0000-0002-2205-7265.

Manuscript received ??; revised ??.

persons participated, but there are also a few studies in which persons with a hearing or visual impairment did the tests. Unless mentioned otherwise in the tables of this paper, participants were both sighted and hearing.

We start this overview with presenting the various building blocks that have been used for words. Some of these building blocks are based on a direct analysis of the input speech signal. Other building blocks based on speech define stimulation patterns that stand for phonemes. A different type of building blocks are letters that are directly presented to various body parts. Finally, there are also building blocks based on Morse code or Braille. These different types of building blocks will be discussed in separate subsections. In the second part of this overview, studies using these building blocks for presenting words are discussed. This will be done in the same order and again in separate subsections as the introduction of the building blocks.

2 TACTUAL SPEECH AND LANGUAGE UNITS

Speech is an acoustic signal and thus the speech signal can be described in terms of physical parameters. But speech can also be described as a series of phonemes and as such phonemes can be considered as building blocks of speech. Language consists of words and the building blocks of words are letters. Letters can have a graphical representation, but they can also be coded as Morse or Braille. In this section, the various building blocks that have been used and tested in tactual devices for speech and language will be presented.

2.1 Tactual Speech Units

There are two fundamentally different approaches to convey tactual speech to a user. In the first approach, the actual speech signal is analysed and somehow converted to a series of vibrations, where different vibrators are linked to different frequency channels. The positioning of the vibrators on the body provides a direct cue to the spectral content of the speech signal. Such devices are generally called “vocoders”. In the second approach, characteristic properties of the speech signal, such as formants or principal components, are somehow transformed to vibratory patterns. An overview of the various studies using one of these approaches is presented in Table 1 and these will be introduced in the sequel of this section.

One of the earliest attempts at conveying speech units by means of a vibratory device was that of Clements et al. [13]. Their device consisted of a square matrix of 12-by-12 vibration motors (also called tactors) applied to the upper thigh. They used two different methods to transform the speech signal into a pattern of vibrations. The two methods differed in that they focused on different aspects of the speech signal. As input speech they used syllables spoken by four different speakers. The task of their two participants consisted of identifying the order in which two syllables that differed in only one phoneme was presented. Phonemes are the smallest units of speech that distinguish one word or part of a word from another. The distinguishing phoneme could either be a vowel or a consonant, but these were tested in separate sessions. It is noteworthy that initial

training took already 20 hours and many additional hours were needed when the stimulus set or the display changed. Even after all these hours of training overall discrimination performance was less than 80%.

Weisenberger et al. [14] tested a 16-channel vibrotactile vocoder applied to the forearm and a 16-electrode electrotactile vocoder applied to the abdomen with three participants. The stimuli were all live-spoken by a female speaker. After many hours of training, participants had to perform several discrimination and identification tasks of syllables that had one distinguishing phoneme. They did not find differences in performance between the two devices. Using a 2-alternative forced choice task (2AFC), accuracy was above 70%, but this dropped to 61% for 8AFC identification of vowels, and to 52% in a 21AFC identification task of initial consonants. In another approach, Weisenberger et al. [15] designed and tested another 30-vibrator device for the forearm based on the principal component analysis (PCA) of the speech signal. In one of their tasks, participants had to identify initial consonants from a set of 9, 10 or 19 syllables of which only the initial consonant differed. For trained participants, performance dropped from 38% correct to 23% if the set was twice as big. For hardly trained participants, performance on the set of 19 consonants was only 12% correct. As overall performance of the trained participants was not too different from the results obtained with the vibrotactile vocoder [14], the authors conclude that PCA may be a promising alternative to a vocoder approach. In a third study, Weisenberger and Percy [16] tested a 7-channel vibrotactile device (the Tactaid VII) that represents the first and second formants of the speech signal. Participants underwent about 80 hours of training and testing. In several experimental sessions, participants had to discriminate or identify vowels, and consonants in various vowel contexts. Their main conclusion was that the vowel context in which a consonant was presented had a large influence; all consonant pairs presented around /æ/ were correctly discriminated more than 90%, whereas this high level was not reached for /i/ or /u/.

Galvin et al. [17] compared performance with two devices, one with 2 tactors (Tactaid II+), the other with 7 (Tactaid VII). The latter one was the same device used by [16]. Their participants were adults with hearing-impairments who used the devices daily during a period of about ten weeks. The 2-tactor device consisted of two tactors on the wrist that responded to the energy level in two different frequency bands. During formal testing, participants were presented with a series of three monosyllabic words differing in one phoneme and they had to decide whether the third word was the same as the first or the second. Performance in this task was above chance level and slightly better with the 2-tactor device than with the 7-tactor device.

Wong et al. [18] mapped articulation location to location on the forearm. In their first experiment, they placed a number of tactors on a sleeve to test how well participants could localize the vibration and how many locations could be distinguished. They found that (only) 2 to 3 tactors could reliably be identified based on location of the vibration. In a subsequent experiment, they converted spoken syllables to vibration patterns on 6 equally-spaced tactors on the forearm. Participants had to discriminate the order in which

TABLE 1
Tactual speech units. The studies in this table use actual speech as input for their devices.

Reference	Location	Participants	Stimuli & Training	Task	Outcome
Clements et al. (1988) [13]	thigh	(1f, 1m)	- 144 vibrators (12 × 12) - 2 methods to represent phonemes based on speech signal - 100 hours training & testing	discriminate syllables (2AFC)	percentage correct: - consonants (73), vowels (84) - spectral method (79), area function method (76)
Weisenberger et al. (1989) [14]	forearm, abdomen	17–23 (1f, 2m)	- array of 16 vibrators or 16 electrodes with input from different frequency channels - vibration patterns representing word pairs differing in 1 phoneme - more than 25 hours testing	discriminate (2AFC) or identify (8, 21AFC) vowels or consonants in word context	- no difference in accuracy between body sites or between articulatory features - 2AFC: above 70% correct - 8AFC of vowels: 61% correct - 21AFC of initial consonants: 52% correct
Weisenberger et al. (1991) [15]	forearm	20–27 (6f)	- 30 vibrators (6 × 5) with input from a PCA of the speech signal - 3 participants with almost a year of daily training	identify consonant	- trained participants: 38%, 39%, 23% correct with sets of 9, 10, 19 consonants - hardly trained participants: 24%, 23%, 12% correct with sets of 9, 10, 19 consonants
Weisenberger & Percy (1995) [16]	forearm	21–26 (6)	7 vibrators with input from different frequency channels together representing 1st and 2nd formants - 80 hours training & testing	discriminate or identify consonants and vowels in varying contexts	- vowel context has significant influence on performance - detailed confusion matrices for different contexts
Galvin et al. (1999) [17]	wrist (c1), sternum or abdomen (c2)	30–77 (8h)	- 2 (c1) or 7 (c2) vibrators conveying frequency and amplitude information from the speech signal - word pairs differing in 1 phoneme	identify stimulus (2AFC)	- performance slightly better with 2 vibrators (c1) - performance not the same for all phoneme contrasts
Wong et al. (2010) [18]	forearm	- 19–27 (2f, 6m) - 18–33 (2f, 2m)	- 5 or 10 equally spaced tactors - 6 equally spaced tactors representing articulation patterns	- identification (5AFC) - rate distance from elbow (scale 0–100) - discriminate 12 consonants presented in pairs (2AFC)	- 81% correct location identification - linear relationship between physical and perceived distance - speech-to-touch method adequate for distinction of place of articulation - method works especially well for fricatives and affricates
Turcott et al. (2018) [19]	forearm(s), upper arm(s)	- 24–54 (2f, 8m)	- 4 displays of 8 equally-spaced tactors - 3 algorithms to convert speech to vibration patterns representing 18 word pairs	- decide whether words differing in one phoneme were same or different	discrimination sensitivity was better for algorithms using Frequency Decomposition and Dominant Spectral Peaks (DSP) than for Autoencoder
Fontana-de-Vargas et al. (2019) [20]	forearm	22–43 (6f, 8m)	- 2 vibrators - patterns related to the vocalization of 24 English phonemes - 100 minutes training	identify phoneme, either individual or in word	- percentage correct: plosives 99%, nasals and approximants 92%, fricatives 70%, diphthongs 95%, vowels 75% - overall phoneme score if tested in words: 68%

For participants age range is given in years (if known) and between brackets the number of participants (f females, m males, h hearing-impaired). If several numbers of participants are itemized, this refers to different experimental conditions or experiments within the paper. xAFC stands for x-alternatives forced choice and the number x in front of it indicates the number of alternatives. PCA stands for principal components analysis. cy stand for condition y, where y is the number of the condition.

two initial consonants were presented to them; these consonants mainly differed in location of articulation. Their main conclusion was that performance was better with fricatives than with plosives.

Turcott et al. [19] tested various algorithms to convert the speech signal into a vibration pattern on the arm. Two of these algorithms coded different characteristics of the speech signal, such as the frequency content (FD) or the

dominant spectral peaks (DSP). A third algorithm (autoencoder or AE) used artificial intelligence to create a representation of the speech signal. Recordings of 18 pairs of words were coded by the algorithms and participants had to indicate whether two words that differed in either zero or one phoneme were the same or different. They found that the use of AE caused lower discrimination sensitivity than the other two algorithms.

A quite recent study using a vibratory device that presents patterns based on the audio wave-form was done by Fontana de Vargas et al. [20]. Their device consisted of only two vibrators. Their mapping from audio signal to vibration pattern was based on various elements related to the vocalization of phonemes. During a total time of only 100 minutes (so less than 2 hours), participants performed self-training sessions to learn subsets of phonemes. In a pre-test in which recognition of individual phonemes was tested, they found correct scores for plosives of 99%, 92% for nasals and approximants, 70% for fricatives, 75% for vowels and 95% for diphthongs. They also tested words and then overall performance for phonemes was 68% correct.

2.2 Tactual Phonemes

As phonemes are the smallest units of speech, it is not surprising that these units have often been taken as a basis for tactual communication devices. In Table 2 the studies that attempt to convey phonemic information are presented. Generally, the relationship with the actual speech signal is quite loose. Instead, the vibratory stimulation patterns are designed in such a way that they will be easy to learn and easy to distinguish.

Dunkelberger et al. [21] constructed a device termed MISSIVE that was able to generate different types of stimulation on the upper arm, namely vibrations, lateral skin stretch and radial squeezes. They defined all English phonemes as different combinations of the possible stimulations. During 100 minutes of training spread over 4 days, participants learned 23 phonemes. Participants could choose their own pacing through the phonemes and they found an average time per phoneme of 3.5 s. Participants rated their confidence in phoneme identification as 3.2 on a scale from 1 to 5. Phonemes without a squeeze (the consonants) were easier to interpret than those with (the vowels). As the focus of this study was on learning tactile words, the authors did not provide separate scores for phoneme identification.

Zhao et al. [22] tested a vibrotactile device consisting of 6 (2-by-3) tactors placed on the dorsal side of the forearm. They defined 9 different phonemes as either one vibration or two sequential vibrations. In one experimental condition (their study 2), participants first learned phonemes before they had to learn words; in another condition (their study 3), the opposite approach was taken: participants first learned words and were later tested on phonemes. In study 2, participants reached an accuracy of about 80% on phoneme identification, whereas in study 3 the accuracy was 91%. So performance on phonemes was better with the top-down approach, in which first words were learned.

Chen et al. [23] defined 13 different phonemes as vibrotactile patterns (24 tactors) on the arm. They compared a guided learning protocol with self-guided learning, both in three sessions on different days. Phonemes were tested in an AXB task, where X was the phoneme pattern they had to recognize and A and B were masking patterns. They found that identification accuracy for phonemes was about 80% in the self-guided group and about 85% in the guided group.

Jiao et al. [24] (also published as part of [26]) presented a tactile sleeve, a vibrotactile device consisting of 24 tactors to be worn around the forearm. Different vibration

patterns represented 39 English phonemes. Although articulatory properties were taken into account in the design of the patterns, another important consideration was the perceptual distinguishability of the patterns. In one of their experiments, participants learned to recognize increasing numbers of phonemes over a period of 7 days, 10 minutes per day. Only if they reached a percentage correct of 90% were they allowed to continue to the word tests on days 8, 9 and 10. Two participants reached this required level on day 7, six on day 8, two on day 9, and the remaining two on day 10. This tactile sleeve was also used in the experiments by Jung et al. [25]. Although from a later date, this study seems to be done earlier than [24]. The main finding was that their four participants were able to learn 10 phonemes in a series of three 10-minute sessions. An extended version of this experiment with additional participants was published as part of [26]. An extensive test of the tactile sleeve was done by Reed and colleagues [27]. Participants were trained with feedback on increasing sets of phonemes, first only consonants, then vowels and finally the total set of 39 phonemes. After they reached a performance criterion of more than 80% correct, which took from 50 to 230 minutes for the various participants, the actual test without feedback began. Averaged over the 10 participants, the recognition rate was 86%.

2.3 Tactual Letters and Digits

Instead of communicating via phoneme conversion to touch as discussed in the previous subsection, one could also use letters as the basic units to be communicated. The current subsection presents an overview of devices that somehow project letters or digits via touch on various body parts (see Table 3). We explicitly exclude numbers presented as sequences of vibration pulses that can be counted, subitized or estimated.

With the so-called Kinotact device of Craig [28], participants could move fiberboard block letters over an array of 10-by-10 photocells that were connected to a 10-by-10 grid of tactors mounted on the back of a chair. The letters covered only a part of the grid, but they could be moved over the whole grid. In several experiments, participants had to identify letters or scrambled letters (abstract patterns consisting of letters of which the columns or rows were interchanged). Performance with the normal or scrambled letters in terms of percentage correct (more than 70% correct after 300 trials with 26 letters) or confusion matrices hardly differed.

An even larger grid of 20-by-20 tactors mounted on the back of a chair was used by Loomis [29]. Participants had to recognize block letters that were randomly presented in 4 different ways. Performance was worst with a static presentation of the letters (34% correct) and best with a static slit behind which a letter moved horizontally (51% correct), and in between with a static letter over which a vertical slit moved (47% correct) and a moving letter (41% correct). Performance of the 3 blind participants was much better than the sighted participants. This is probably not surprising as only the blind participants had a training of 100 hours before the actual experiment. In a follow-up experiment, it was shown that performance with a diagonal static slit

TABLE 2
Tactual phonemes. The studies in this table define phonemes by different stimulation patterns.

Reference	Location	Participants	Stimuli & Training	Task	Outcome
Dunkelberger et al. (2018) [21]	upper arm	19–30 (6f, 4m)	- vibrotactor band, radial squeeze band, haptic rocker - patterns representing 23 English phonemes - 100 minutes training	- rate confidence of phoneme identification	- confidence: 3.2 on scale from 1–5 - average self-paced time per phoneme: 3.5 s - squeeze makes phonemes harder to identify
Zhao et al. (2018) [22]	forearm	20–50 - (8f, 3m) - (4f, 5m)	- 6 tactors - layout similar to Braille representing 9 phonemes - various training protocols	identify phoneme	percentage correct: - 80% after training on phonemes - 91% after training on words
Chen et al. (2018) [23]	arm	21–43 (6f, 13m)	- 3 displays of 8 (4 × 2) vibrators - patterns representing 13 phonemes - 2 learning methods - 65 minutes training	identify phoneme	accuracy in self-guided learning group: 80%, in guided learning group: 85%
Jiao et al. (2018) [24]	forearm	- 18–26 (6f, 6m)	- 24 tactors (4 × 6) - patterns representing 39 phonemes - 100 minutes training/testing	identify phoneme	number of 10-minute sessions to reach a level of 90% correct: 7 (2 participants), 8 (6 participants), 9 (2 participants), 10 (2 participants)
Jung et al. (2019) [25], Tan et al. (2020) [26]	forearm	20–30 (4)	- 24 tactors (4 × 6) - patterns representing 10 phonemes - 30 minutes training/testing	identify phoneme	30 minutes sufficient to learn 10 phonemes
Reed et al. (2019) [27]	forearm	19–32 (7f, 3m)	- 24 tactors (4 × 6) - 39 English phonemes represented by different activation patterns - 50–230 minutes training	identify phoneme	mean recognition rate of 86% correct

For participants age range is given in years (if known) and between brackets the number of participants (f females, m males, h hearing-impaired). If several numbers of participants are itemized, this refers to different experimental conditions or experiments within the paper. xAFC stands for x-alternatives forced choice and the number x in front of it indicates the number of alternatives. PCA stands for principal components analysis. cy stand for condition y, where y is the number of the condition.

behind which a letter moved was only slightly worse than with a vertical static slit.

The device of Saida et al. [30] consisted of a grid of 10-by-10 tactors placed against the abdomen. They presented 46 Katakana characters (Japanese letters) in three different modes: static (stationary presentation), moving (characters pass horizontally over the back) and tracing. In the tracing mode, the character was presented as if it was being written. All characters had to be recognized once in each of the three conditions. Performance with the traced characters (95% correct) was much better than with the moving (39% correct) or static (27% correct) presentations. Interestingly, they did not find a difference in performance between their blind and sighted participants.

Yanagida et al. [31] aimed at creating a relatively simple vibrotactile device with only 9 (3-by-3) tactors embedded in a chair capable of representing 26 letters and 10 digits. They designed a clever way to translate the tracing of a letter to a pattern of vibrations by projecting the letters and digits on a grid with unequal partitions. Without any training, their participants were able to recognize 87% of the letters and 88% of the digits.

Kim et al. [32] placed an array of 5-by-5 tactors on top of the foot. They compared two different modes of sequential stimulation. In both cases, the tactors were stimulated in

the same order as would be the case in handwriting. In the conventional mode, the tactors vibrated one by one, whereas in the advanced mode, there was always some overlap in stimulation between subsequent tactors, causing a more continuous flow of stimulation. After training, performance in the conventional mode was 60% correct letter recognition, whereas in the advanced mode a level of 87% correct was reached, suggesting that continuity of motion is an important feature for letter recognition. The advanced mode (called improved handwriting mode) was also tested with an 8-by-6 array of tactors on the back by Wu and colleagues [33]. They compared this mode with a horizontal slit moving over a letter (similar to the vertical slit of [29]) and their own improved mode, where vertices of the letters were marked with higher vibration intensities. Participants had to identify 10 letters presented several times with the three vibration modes. Not surprisingly, worst performance (48% correct) was obtained with the scanning mode with the slit. The advanced mode gave much better results (77% correct), but their new mode gave the best results (82% correct).

The focus of the research of Arnold and Auvray [34] was to determine how well training of discriminating vibratory letters on one body part transferred to another body part. Participants had to discriminate 8 capital letters presented

TABLE 3
Tactual letters and digits. The studies in this table create stimulation patterns that simulate or define letters and/or digits.

Reference	Location	Participants	Stimuli & Training	Task	Outcome
Craig (1973) [28]	back	- (4) - (6) - (5)	- 100 tactors (10 × 10) - 6 (9, 26) vibration patterns representing normal or scrambled letters - 10 minutes warm-up - active pattern exploration	identify 6, 9, or 26 letters	- hardly any decrement for scrambled letter patterns - good transfer between normal and scrambled letter patterns - confusion matrices similar for the 2 types of patterns
Loomis (1974) [29]	back	20–35 (2f, 5m; 3b, 4s)	- 400 tactors (20 × 20) in chair - 26 block letters patterns - 5 modes of presentation - blind participants: 100 hours training	identify letter	- individual performances quite different - percentage correct: static letters (34), static vertical slit, moving letters (51)
Saida et al. (1982) [30]	abdomen	30–36 (4b, 4s)	- 100 vibrators (10 × 10) - 46 Katakana characters - 3 modes of presentation: static, moving, tracing	identify Katakana characters	- percentage correct: static (27), moving (39), tracing (95) - no difference between performance of blind and sighted participants
Yanagida et al. (2004) [31]	back	22–39 (2f, 8m)	- 9 tactors (3 × 3) embedded in chair - vibration patterns representing 26 letters & 10 digits - no training	identify letter and/or digit	- percentage correct: digits (88), letters (87), mixed (86) - use of stroke-related information is promising
Kim et al. (2006) [32]	foot	26–31 (10m)	- 25 vibrators (5 × 5) - 2 modes of sequential stimulation: single or pair of motors	identify letter	percentage correct (without/with training): - single motor (32, 60) - two simultaneous motors (68, 87)
Wu et al. (2012) [33]	back	25–30 (15)	- 48 vibrators (8 × 6) - 10 letters - 3 modes of presentation: scanning, handwriting, tracing	identify letter	percentage correct: - scanning (48) - handwriting (78) - tracing (82)
Arnold & Auvray (2014) [34]	belly, shin, thigh	18–47 (26f, 19m)	- 9 vibrators (3 × 3) - 8 vibration patterns representing drawing letters	identify letter (8AFC)	- significant improvement over sessions - training of one body site transfers to other sites
Liao et al. (2016) [35]	wrist	- 21–29 (13f, 11m) - 21–25 (12f, 12m)	- 4 (2 × 2) tactors - vibration patterns representing 26 letters & 10 digits - 15 minutes training	identify alphanumeric characters	- identification unistroke: letters: 71%, digits: 79% - identification multistroke: letters: 86%, digits: 89%
Janidarmian et al. (2018, 2019) [36], [37]	back	18–46 (5f, 5m)	- 9 vibrators (3 × 3) - vibration patterns representing 26 letters & 10 digits - standard or personalized - training: 108 trials	identify alphanumeric characters	percentage correct: - standard pattern (71), - personalized pattern (87)
Cunha & Nohama (2015) [38]	abdomen	(13)	- patterns of gas jets representing 19 letters and 3 geometric figures - 30 minutes training	identify 11 letters	- accuracy depends on “plotting” speed (highest: 68% with 40 mm/s)

For participants age range is given in years (if known) and between brackets the number of participants (f females, m males, b blind, s sighted). When not mentioned, participants were sighted and hearing. If several numbers of participants are itemized, this refers to different experimental conditions or experiments within the paper. 8AFC stands for 8-alternatives forced choice.

on a 3-by-3 grid of tactors as if they were handwritten. Training on either belly, shin, or thigh significantly improved performance both in percentage correct and in speed, but notably, a post-training session showed that this learning transferred completely to the other body parts.

Liao et al. [35] tested a vibrotactile wrist device consisting of 4 (2-by-2) tactors. They generated vibration patterns based on EdgeWrite [39] representing alphanumeric characters. The vibration patterns for the various characters consisted of 2 to 6 vibrations. In their “unistroke” condition they found that especially the 2- and 3-vibration patterns had a high recognition rate. Therefore, in their “multistroke” condition, they split the patterns consisting of 4 or more vibrations into a sequence of 2- or 3-vibrations such that

they became easier to recognize. Recognition performance improved to 86% correct for letters and 89% correct for digits after only about 15 minutes of training.

Janidarmian et al. [36], [37] presented alphanumeric characters on a 3-by-3 grid of tactors on the back. The order of the stimulation of the various tactors was determined by how a lefthanded experimenter would write them. After a training session in which all characters were presented three times, identification performance was 71%. Part of the confusions were due to different handwriting habits of the participants. In the second session, participants could change the order of stimulation of the tactors in how they would write the various characters. With these personalized patterns, identification performance increased to 87%. Note,

however, that this increase could partly be due to an increased familiarity with the device as this second condition was always measured after the first one.

A completely different approach to convey letters to a body part was taken by Cunha and Nohama [38]. With their fluxtactile stimulator they could write letters by means of gas jets on the abdomen of participants. After a limited training period of 30 minutes, their participants reached a performance level of 68% correct letter recognition. The advantage of this method is that it produces continuous stimulation instead of discrete stimulation as with the vibrotactile devices. However, the hardware needed is quite bulky and the letters have to be written on naked skin, making this device less suitable for daily life use.

2.4 Tactual Morse and Braille-like Codes

Another way to code letters is via Morse code, where dots and dashes consist of short and long vibrations, respectively. Ranjbar et al. [40] asked proficient auditory Morse users to test their device on both the wrist and the abdomen. Performance was quite good (93% and 86% correct, respectively) and did not differ significantly between the two body sites. However, two participants did not feel any vibrations on their abdomen and these were left out of the analyses. Seim et al. [41] studied passive haptic learning (PHL) of Morse code letters while playing a distractive game. Mode of transfer of the Morse code was via bone conduction on a Google glass. Performance of the PHL group was significantly better than the control group. After four hours of playing the game, the PHL group was able to write the Morse codes of letters with 98% accuracy and recognize the Morse codes with 86% accuracy. In a follow-up study on PHL, Seim et al. [42] presented 10 Morse code letters via a smartwatch and at the same time participants heard the letter via headphones. They found again that passive haptic learning of Morse code is well possible. Pescara et al. [43] repeated the experiments of [41] and [42] using a wrist device. In additional conditions, they precluded any feedback during the tests. Their clear outcome was that although a little learning also occurred during passive learning, active learning gave much better results.

In the study of Plaisier et al. [44], participants unfamiliar with Morse code had to learn letters in groups of three. Dots were presented as short vibrations to their left arm, dashes as longer vibrations to their right arm. After a practice series followed a test series. Only if all three letters were correctly identified during this test phase there followed a new practice series with three additional letters. The duration of this training and testing session was only 30 minutes. Four participants managed to learn 25 letters during this time, one participant learned 15 letters and the other five somewhere in between. It is noteworthy that participants never saw or heard a visual or auditory representation of Morse code.

Letters can also be coded as Braille characters and that was the approach taken by Schätzle et al. [45] with their vibrotactile wristband called VibroTac. Braille characters consist of 6 dots and therefore their device consisted of 6 equally-spaced tactors around the wrist. In this study, they tested 8 different methods of mapping Braille characters to

vibrations patterns. A clear outcome was that performance with static patterns (in their terms parallel patterns) was much worse than with sequential patterns. Performance was best if not only the raised dots were vibrating but also the unraised dots with lower intensity.

2.5 Tactual Speech and Language Units: Summary and Conclusions

The impression that is left when considering the majority of the devices based on actual speech input is that performance is relatively low, often just above chance level, even with many hours of training. The only positive exception seems to be the recent study of Fontana-de-Vargas et al. [20]. The devices that designed vibrotactile patterns for different phonemes are in general more successful, although many of them only used a limited set of phonemes. Especially the studies that used the tactile sleeve approach and a full set of English phonemes [24], [27] are quite promising, both in terms of performance as well as in required training time. Interestingly, several devices presenting letters in some way to a body part did not need training. A disadvantage of these devices is probably that they will be slow in conveying longer messages. The same holds for devices based on Morse code or Braille. An advantage of these latter devices is that they can be rather simple, small and easy to put on.

There are also some more general conclusions that can be drawn from the results presented in this section. First, performance with static presentation of vibratory patterns is much worse than with dynamic patterns [29], [30], [45]. Second, devices with many tactors (25–400) do not really outperform devices with only a small number of tactors (1–9). This strongly suggests an advantage for devices with fewer tactors as these will be much cheaper and lighter. Third, good performance is often possible with hardly any training, especially when letters are presented directly to a body part [31]. Most probably, with a few hours of additional training, performance close to perfect could be reached. Fourth, subtle changes in transfer or presentation mode, such as the use of a non-standard grid [31], overlapping stimulation [32] or marking vertices [33] may cause a substantial increase in performance. Finally, there are some indications that what has been learned at some body part may transfer to other body parts [34], but this is based on just a small experiment and needs to be investigated more extensively.

3 TACTUAL WORDS

Creating word building blocks by means of a tactile device is only the first step. Ultimately, one wants to transmit words, either in isolation or in sentences. In this section, we give an overview of devices that have been tested to convey hands-free tactual words. The order of presentation will be the same as in the previous section. We start with studies on devices like vocoders that are based on a direct analysis of the speech signal (see Table 5).

3.1 Tactual Words based on Speech Signal Characteristics

Brooks and Frost [46] made a major effort in teaching two participants words via a vocoder on their forearm.

TABLE 4

Tactual letters presented via Morse or Braille-like codes. The studies in this table present devices that produce Morse or Braille-like codes.

Reference	Location	Participants	Stimuli & Training	Task	Outcome
Ranjbar et al. (2016) [40]	abdomen, wrist	40–85 (14m)	- 1 tactor - Morse patterns of 29 letters	identify letter	- % correct: abdomen (86), wrist (93) - some participants had 100% correct
Seim et al. (2016) [41]	head	18–25 (5f, 7m)	- Google glass - 4 hours of passive haptic learning (PHL) of Morse code	reproduce & recognize letters	- 98% correct Morse code reproduction - 83% correct letter recognition
Seim et al. (2018) [42]	wrist	(1f, 5m)	- smartwatch - passive learning of 10 Morse code letters during 8 or 16 min/word	reproduce Morse code	- significant improvement between pre- and post-tests - passive learning possible
Pescara et al. (2019) [43]	wrist	18–28 (17f, 33m)	- 1 tactor - learn 10 random Morse codes - conditions based on [41], [42]	- identify digit - reproduce Morse code	- passive learning in [41] and [42] mainly due to active tests with feedback
Plaisier et al. (2020) [44]	forearms	20–26 (4f, 6m)	- 2 tactors (left: dots, right: dashes) - 30 minutes learning of Morse code letters in groups of 3	identify letter	1 participant learned 15 letters, 4 learned 25 letters, the others in between
Schätzle et al. (2017) [45]	wrist	- 22–34 (1f, 3m) - 16–48 (7f, 11m)	- 6 vibrators in linear array representing Braille characters - 8 or 3 mapping methods - few trials training	- draw Braille character	- sequential methods: better performance - better performance with 2-dot than with more than 2 dots Braille characters - best performance if unraised dots are also represented (97% correct)

For participants age range is given in years (if known) and between brackets the number of participants (f females, m males, b blind, s sighted). When not mentioned, participants were sighted and hearing. If several numbers of participants are itemized, this refers to different experimental conditions or experiments within the paper. 8AFC stands for 8-alternatives forced choice.

TABLE 5

Tactual words - stimulation based on speech signal characteristics

Reference	Location	Participants	Stimuli & Training	Task	Outcome
Brooks & Frost (1983) [46]	forearm	(2f)	- tactile vocoder - 16 spectral filter channels activating 16 solenoids coding spoken language - 40–55 hours training in sessions of 20 minutes	learn words	- participant 1: 70 words in 81 sessions - participant 2: 150 words in 110 sessions - rapid transfer to words spoken by other person
Brooks & Frost (1985) [47]	forearm	(1f) participant 2 of [46]	- tactile vocoder - 16 spectral filter channels activating 16 solenoids coding spoken language - 26 hours training in sessions of 20 minutes	learn additional 100 words	- 26 hours needed to learn the 100 additional words - learning rate 3.9 words/hour
Lynch et al. (1988) [48]	abdomen	20, 32 (1fd, 1md)	- array of 16 electrodes representing different frequency bands - more than 41 hours training on 50 words	recognize word	overall 42% correct of 11,324 trials
Turcott et al. (2018) [19]	forearm(s), upper arm(s)	22–44 (4f, 12m)	- 2 or 4 displays of 8 tactors - 2 algorithms to convert speech to vibration patterns representing 24 words	play a word learning game	in game learning the Phonemic algorithm yielded better scores than the Dominant Spectral Peak (DSP) algorithm
Fontana-de-Vargas et al. (2019) [20]	forearm	22–43 (6f, 8m)	- 2 vibrators - patterns resembling 24 English phonemes - 100 minutes training	- recognize words - multiple choice (50 words)/open test	accuracy: - multiple choice: words 94.4% - open test: words 45%

For participants age range is given in years and between brackets the number of participants (f females, m males); d indicates deaf participants.

The participants underwent 40 to 50 hours of training in sessions of 20 minutes. The words were actually spoken and subsequently transmitted to the forearm via a vocoder. One of the participants managed to learn 70 words in 81 sessions,

whereas the other person learned 150 words in 110 sessions. Both participants were able to rapidly understand words spoken by another person. In [47] these authors continued their study with the participant who had learned 150 words,

by teaching this person an additional 100 words. To reach a level of 250 words took an additional 26 hours. So although understanding words via a vocoder is possible, the learning rate was quite slow, namely 3.9 words per hour.

Tactile communication devices are often meant for persons who are either visually or auditory impaired. It is therefore also important to test the devices with the target group. Lynch and colleagues [48] trained two deaf persons on word recognition via an electrotactile device on their abdomen. The 16 electrodes represented different frequency bands of the speech signal. More than 40 hours were required to teach the two participants 50 words. Overall performance over more than 11,000 trials was 42% correct. So although it is possible to learn words via this tactile device, it seems rather inefficient to invest so many hours on learning to recognize words presented to the abdomen.

Turcott et al. [19] designed a computer game in which different levels could be reached by recognizing 2- and 3-phoneme words presented as vibrotactile patterns using two different algorithms for different groups of participants. With both algorithms the participants performed well above chance, but performance in terms of both recognition rate and retention was better with an algorithm based on phoneme-inspired patterns.

Fontana de Vargas et al. [20] used their two-vibrator device not only to present individual phonemes (see Table 1), but also to present 50 words. Word length ranged from one to six phonemes. The temporal spacing of the phonemes of a word was controlled by the participants. After a total training time of 100 minutes (this includes the training time for phonemes described above), their participants were able to reach a 94% correct score on a multiple choice test of 50 words. If the test was open, so no answer choices were provided, performance was still 45% correct.

3.2 Tactual Words based on Phonemes

In Table 6 we present and discuss hands-free devices that convey words as sequences of phonemes. These studies were already introduced above in Table 2 as they also report results on the tactile perception of phonemes. Here we focus on the parts of those studies that describe the results on word perception.

The device constructed by Dunkelberger et al. [21] was used to generate stimulation patterns consisting of vibrations, lateral skin stretch and radial squeezes. Different patterns were assigned to different English phonemes. During a total training time of 100 minutes, participants were familiarized with 150 1- to 6-letter words consisting of these phonemes. In a final test they had to recognize 50 of the words they had learned. The average score in this final test was 87% correct. The identification confidence score of words was 4.1 on a scale from 1 to 5, which was higher than that for phonemes (score 3.2). The authors conclude that their results are promising for the creation of a tactile language.

The 9 different phonemes that could be presented with the vibrotactile device of Zhao et al. [22] were used to create two sets of 10 words. In their study 2, the first set was used for training, whereas the second set was used for a final test. Only during training participants received feedback

and could use a reference sheet explaining the tactile codes. Performance on new words identification was significantly better if the mapping of vibration pattern to phoneme was somewhat based on articulation location instead of random (46% versus 27%). In their study 3, a top-down approach was tested, in which directly words were learned instead of phonemes. Their main conclusion was that performance in terms of word identification accuracies was similar in the two studies, but training time was shorter for the top-down approach.

Chen et al. [23] compared the efficiency of different approaches to learn 100 2- and 3-phoneme words consisting of 13 different phonemes presented as vibrotactile patterns (24 tactors) on their arm. The self-guided group had more freedom to explore the words at their own space. During the first test, the guided learning group showed better performance, but at all later stages there were no significant differences. In the final test, accuracy was 86% for the guided group and 72% for the self-guided group, but this difference was apparently not significant. However, in the guided learning group significantly more participants reached a level of more than 90% correct word recognition after a total of only 65 minutes of training and testing. Performance on the phoneme identification test was a good predictor of the word score, explaining 70% of the variance.

Jiao et al. [24] used two different approaches to learn 2- and 3-phoneme words with their tactile sleeve. In one approach, participants had to first learn all phonemes over a period of 7 days and reach an accuracy of 90% before they were allowed to continue to the word phase. All participants eventually managed to reach the required level after 7 to 10 days, but their scores on word identification varied substantially. Still, 6 of their participants reach a score higher than 80% on an identification test of 100 English words after a total learning period of only 100 minutes. In the other approach, participants started to learn words right from day 1. Over the days, the words were built from increasing numbers of different phonemes. In this experiment, the differences between participants were even larger. Two of the participants managed to learn 80 words, whereas for the other participants the learning curve seemed to flatten after day 8 and remain at a level of 50 words. This study was also published as experiment 2 of [26].

Tan et al. [26] further extended the work with the tactile sleeve. In their first experiment, an extension of [25], 10 participants had to learn 10 phonemes and 51 2- or 3-phoneme words. All participants were able to learn the words within an hour with very few errors. Their second experiment was also published as [24], so see above. In a third experiment, the number of words to be learned was extended to up to 500. Over a period of a few weeks (only a few hours per week, with a total maximum of 8 hours) participants advanced from level to level, so within the experimental time, some participants reached higher levels than others. The best participant managed to learn 469 words, while the worst performing participant still learned 131 words; however, most participants learned at least 300 words. The authors conclude that their results demonstrate the feasibility of their tactile sleeve as a tactile communication device.

TABLE 6
Tactual words consisting of phonemes

Reference	Location	Participants	Stimuli & Training	Task	Outcome
Dunkelberger et al. (2018) [21]	upper arm	19–30 (6f, 4m)	- vibrotactor band, radial squeeze band, haptic rocker - patterns representing 23 English phonemes - 100 minutes training	- recognize words - multiple choice (50 words) - rate identification confidence (1–5)	- word recognition accuracy: 87% - identification confidence: 4.1
Zhao et al. (2018) [22]	forearm	20–50 - (8f, 3m) - (4f, 5m)	- 6 tactors - layout similar to Braille representing 9 phonemes - various training protocols	recognize 2- and 3-phoneme words	- training time shorter when trained on words - performance better if phoneme patterns are based on articulation
Chen et al. (2018) [23]	arm	21–43 (6f, 13m)	- 3 displays of 8 (4 × 2) vibrators - patterns representing 13 phonemes used to form 100 1–3-phoneme words - 2 learning methods - 65 minutes training	type the word	- initially performance better with guided learning than with self-guided learning - percentage correct: guided learning 86, self-guided learning 72 - performance on phoneme test good predictor for word score
Jiao et al. (2018) [24]	forearm	- 18–26 (6f, 6m) - 19–39 (6f, 6m)	- 24 tactors (4 × 6) - patterns representing 39 phonemes used to form 100 1–3-phoneme words - 2 learning methods - 100 minutes training/testing	identify 2- and 3-phoneme words	- learning method: phoneme-based gives better results than word-based - large differences between participants - 8 of 24 participants learned 80 words in 100 minutes
Tan et al. (2020) [26]	forearm	18–39 (27f, 28m)	- 24 tactors (4 × 6) - patterns representing 39 phonemes used to form 500 1–3-phoneme words - max 8 hours of training	identify word	- near perfect performance for learning 51 words in 1 hour - see [24] for conclusions from experiment 2 - 65% correct word identification

For participants age range is given in years and between brackets the number of participants (f females, m males). If several numbers of participants are itemized, this refers to different experimental conditions or experiments within the paper.

3.3 Tactual Words based on Letters

Of all the studies in Table 3 that tested devices for the presentation of letters, digits or characters, only Janidarmian et al. [37] made the follow-up step to also test word recognition (see Table 7). Their personalized way to present letters on the back via 3-by-3 tactors as described in the previous section [36], [37] was also used to test word recognition [37]. Without any additional training, participants were asked to recognize 20 2- to 4-letters verbs. Recognition performance was 91%.

3.4 Tactual Words based on Morse Code or Arbitrary Patterns

In Table 8 we present devices that use Morse code or arbitrary patterns to construct words.

A few experiments with a rather limited set of words were done by Velazquez and colleagues [49], [50], [51] using a tactile display with 4 tactors stimulating the foot sole. They constructed words by a series of simultaneous vibrations of short and long durations of all 4 tactors [51]. During a short training session, participants had to learn the vibratory patterns of 4 or 5 words, the identification of which was tested shortly after. They also tested “sentences”, sequences of 2, 3, or 4 of the words. Recognizing 1 word out of a set of 4 gave a 84% correct score, and sentences consisting of 4 words were 66% correct. They found this a promising way of constructing a tactile language for human computer

interaction. However, the set of words they can create with this device seems quite limited.

Ranjbar et al. [40] used vibratory Morse code to present 15 Swedish riding terms via one tactor to either the wrist or the abdomen of experienced auditory Morse code users. As in the auditory Morse codes, dots and dashes were distinguished by the duration of stimulation (dashes three times as long as the dots). Performance on the wrist (92% correct) was much better than on the abdomen (70% correct). They concluded that their device is suitable for instructing persons with deafblindness riding on a horse, as was their intended application. Also Walker and Reed [52] used vibratory Morse code for presenting sequences of alphanumeric characters. Dots and dashes were either given both to the same arm, or dots were given to the left arm and dashes to the right arm. In the two-arm case, dashes were either three times as long as dots, or of equal duration. In an additional equal-duration condition, both arms had 3 tactors, so that the different dots or dashes of a character stimulated slightly different locations. They did not actually test recognition of words, but sequences of 3 characters composed out of a set of 12 characters. The main outcome was that performance with the two-arm set-up with unequal dot and dash durations was significantly better than performance with the one-arm set-up. Due to these findings, Plaisier et al. [44] also presented vibrations belonging to dots and dashes to different arms. In their experimental protocol, participants had to learn as many letters as possible within

TABLE 7
Tactual words conveyed via letters

Reference	Location	Participants	Stimuli & Training	Task	Outcome
Janidarmian et al. (2019) [37]	back	(10)	- 9 vibrators (3 × 3) - vibration patterns representing 20 2–4 letter verbs - no training	recognize word	percentage correct: 91

The number between brackets in the participants column gives the number of participants.

TABLE 8
Tactual words - language conveyed via Morse code or arbitrary patterns

Reference	Location	Participants	Stimuli & Training	Task	Outcome
Velásquez et al. (2011) [49]	foot	19–23 (6m, 14m)	- 4 (1–2–1) vibrators - 5 temporal patterns representing words	recognize word (5AFC)	overall 75% correct
Velásquez & Pissaloux (2014, 2018) [50], [51]	foot	18–24 (4f, 16m)	- 4 (1–2–1) vibrators - 4 temporal patterns representing words	- recognize word (4AFC) or 2,3,4-word sequence	average recognition performance of sequences: 1 word 84%, 2 words 77%, 3 words 77%, 4 words 66%
Ranjbar et al. (2016) [40]	abdomen, wrist	40–85 (14m)	- 1 tactor - Morse patterns of 15 Swedish riding terms	identify word	- percentage correct: abdomen 70%, wrist 92% - some participants had 100% correct
Walker & Reed (2018) [52]	forearm	20–30 (2f, 6m)	- 12 Morse vibration patterns on one or two arms - dash duration equal/unequal to dot duration	identify sets of 3 Morse characters	- fewer errors with two-arm presentation - no significant influence of dash duration
Plaisier et al. (2020) [44]	forearms	20–26 (4f, 6m)	- 2 tactors (left: dots, right: dashes) - 30 minutes of learning letters in Morse - 20 minutes word testing	identify 2, 3, 4, 5-letter words	correct rates decreased rapidly with word length (from 75% for 2-letter words to 8% for 5-letter words)

For participants age range is given in years and between brackets the number of participants (f females, m males). If several numbers of participants are itemized, this refers to different experimental conditions or experiments within the paper. xAFC stands for x-alternatives forced choice, where x is the number of alternatives.

30 minutes. After this period followed a 20-minute test on words consisting of the phonemes learned. The number of phonemes learned differed between participants, so also the word sets on which they were tested differed. The main finding was that performance decreased rapidly with word length, from 75% correct for 2-letter words to only 8% for 5-letter words.

3.5 Tactual Words: Summary and conclusions

The studies in this section show that staying close to the actual speech signal is not necessarily the best way to proceed. Especially the three old studies in Table 5 show that even with numerous hours of training, word recognition success is quite limited. Only the very recent study in this table [20] gives promising results. The studies in Table 6 that defined vibrotactile patterns for phonemes are all recent and although sometimes limited in scope, show positive results after just a relatively short training period. The most advanced device is the tactile sleeve of Tan, Reed and colleagues [26]. In maximal 8 hours their participants were able to learn up to 500 words. With 500 words one can convey already quite complicated messages and experienced users will be able to easily extend their vocabulary. That makes the tactile sleeve an interesting device for persons with deafblindness.

Words consisting of letters “written” on the back [37] were even without training almost perfectly recognized, which is a real advantage compared to the other devices. A possible disadvantage of this method is that it will be slow to convey longer messages, but for short messages that needs not be a limitation. Finally, Plaisier and colleagues [44] showed that naive participants could learn some basics of vibrotactile Morse code in just half an hour. This time was not sufficient for conveying whole words, but it is not unlikely that with a little additional training, also words could be recognized. The advantage of this method would be that it is cheap and light.

4 GENERAL CONCLUSION

Taking all these studies together, the conclusion seems justified that best successes were obtained with relatively small numbers of vibrators on the back, or with a sleeve with up to 24 vibrators on the forearm. It is clearly possible to recognize vibratory letters on the back, and even to recognize words composed of such letters. The advantage of using letters is that most participants were already familiar with the pattern (i.e. the letter) itself. Learning arbitrary patterns was also possible, but it still remains to be seen how well that is possible for larger numbers of patterns (above 12).

Another clear conclusion is that static presentation of vibrotactile patterns is not the way to go. All studies that compared performance on both static and dynamic patterns showed a very substantial difference in favour of dynamic patterns [29], [30], [45], [53], [54]. Also some other studies that only used static patterns showed low performance even after many hours of training [10], [55], [56].

Most of the studies tested blindfolded sighted persons. Only a few studies worked with blind or deaf individuals. Therefore the relevance of these findings for individuals with blindness or deafblindness remains to be seen, although it seems a safe hypothesis that their success rates will be similar or higher. It should be noted, however, that although persons with deafblindness are often interested to use devices that make their life easier, even learning to use relatively simple devices will cost them a lot energy [personal communication]. Therefore, it seems unlikely that they are willing to invest many hours of training to learn to use a new device or learn a new language, unless it will be clear from the start that it will be of great value for them. That is a challenge that should be kept in mind.

One concern is that many of the studies seem to result from small student projects without a proper follow-up. However, many of the relatively successful studies are quite recent, so follow-ups may still be on the way. Thus, we may hope that this boost of recent interest in tactile communication devices may eventually result in a number of really usable devices. Notably the research with the tactile sleeve [26], [27] went already beyond the pilot stage, as their participants were able to learn a substantial number of words within a reasonable amount of time.

The question arises what further steps need to be taken when developing the next generation prototypes. For most of the devices in this survey, the next step needs to be to increase the number of speech units, phonemes or letters, so that a full vocabulary can be built from the building blocks. A huge subsequent step will be to form sentences. Then speed of presentation may become a bottleneck. In all cases, one needs to consider what the main purpose of the device will be. For conveying only short messages less ambitious goals can be set than if one wants to use the device for interaction with another person.

Another consideration should be the intended circumstances in which to use the device. If it needs to be available in all situations, it also needs to be portable and thus light and small. If the intention is to only use it in an indoor setting, the device may be much more bulky and connected to a computer. Clearly, there is not just one way to go and different prototypes may serve different goals.

Another issue that needs to be considered is the type of input. Is there on-the-fly translation from speech to tactual speech possible? How easily can spoken language be translated in (near) real time to the tactual signal? And if letters are used as building blocks, what kind of additional input device is needed? These issues do not have to be solved right away, but clearly they are relevant when designing a usable product.

A final challenge is that of two-way communication. The devices discussed in this survey all communicate speech or language to a user of the device. For actual communication, the user should also be able to respond or react. When

attending a meeting or a conference and just following the speaker this might not be an issue, but in person-to-person communication this certainly is. Thus it will depend on the environmental situation and the severity of the sensory loss of the user whether the device is suitable for two-way communication. Certainly if the target population is that of users with deafblindness, this is a challenge that also needs to be solved. However, that challenge lies well beyond the scope of the current survey.

ACKNOWLEDGMENTS

The authors have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 780814, project SUITCEYES. The authors also want to thank two anonymous reviewers who gave very valuable advice on the organization of this paper.

REFERENCES

- [1] A. Wilska, "On the vibrational sensitivity in different regions of the body surface," *Acta Physiologica Scandinavica*, vol. 31, pp. 285–289, 1954.
- [2] S. Weinstein, "Intensive and extensive aspects of tactile sensitivity as a function of body part, sex, and laterality," in *The skin senses*, D. Kenshalo, Ed. Springfield, IL: Thomas, 1968, pp. 195–222.
- [3] L. A. Jones and N. B. Sarter, "Tactile displays: Guidance for their design and application," *Human Factors*, vol. 50, no. 1, pp. 90–111, 2008.
- [4] I. Karuei, K. E. MacLean, Z. Foley-Fisher, R. Mackenzie, S. Koch, and M. El-Zohairy, "Detecting vibrations across the body in mobile contexts," in *Proceedings of ACM Conference on Human Factors in Computing Systems (CHI '11)*, 2011, pp. 3267–3276.
- [5] R. W. Cholewiak and A. A. Collins, "Vibrotactile localization on the arm: effects of place, space, and age," *Perception & Psychophysics*, vol. 65, no. 7, pp. 1058–1077, 2003.
- [6] D. A. Mahns, N. M. Perkins, V. Sahai, L. Robinson, and M. J. Rowe, "Vibrotactile frequency discrimination in human hairy skin," *Journal of Neurophysiology*, vol. 95, no. 3, pp. 1442–1450, 2006.
- [7] R. Hoffmann, V. V. Valgeirsdóttir, O. I. Jóhannesson, R. Unnthorsson, and A. Kristjánsson, "Measuring relative vibrotactile spatial acuity: effects of factor type, anchor points and tactile anisotropy," *Experimental Brain Research*, vol. 236, no. 12, pp. 3405–3416, 2018.
- [8] J. C. Craig, "Vibrotactile difference thresholds for intensity and the effect of a masking stimulus," *Perception & Psychophysics*, vol. 15, no. 1, pp. 123–127, 1974.
- [9] P. Eskildsen, A. Morris, C. C. Collins, and P. Bach-y-Rita, "Simultaneous and successive cutaneous two-point thresholds for vibration," *Psychonomic Science*, vol. 14, no. 4, pp. 146–147, 1969.
- [10] R. W. Cholewiak and A. A. Collins, "Vibrotactile pattern discrimination and communality at several body sites," *Perception & Psychophysics*, vol. 57, no. 5, pp. 724–737, 1995.
- [11] S. D. Novich and D. M. Eagleman, "Using space and time to encode vibrotactile information: toward an estimate of the skin's achievable throughput," *Experimental Brain Research*, vol. 233, no. 10, pp. 2777–2788, 2015.
- [12] O. I. Jóhannesson, R. Hoffmann, V. V. Valgeirsdóttir, R. Unnthorsson, A. Moldoveanu, and A. Kristjánsson, "Relative vibrotactile spatial acuity of the torso," *Experimental Brain Research*, vol. 235, no. 11, pp. 3505–3515, 2017.
- [13] M. A. Clements, L. D. Braid, and N. I. Durlach, "Tactile communication of speech: comparison of two computer-based displays," *Journal of Rehabilitation Research and Development*, vol. 25, no. 4, pp. 25–44, 1988.
- [14] J. M. Weisenberger, S. M. Broadstone, and F. A. Saunders, "Evaluation of two multichannel tactile aids for the hearing impaired," *Journal of the Acoustical Society of America*, vol. 86, no. 5, pp. 1764–1775, 1989.
- [15] J. M. Weisenberger, J. C. Craig, and G. D. Abbott, "Evaluation of a principal-components tactile aid for the hearing-impaired," *Journal of the Acoustical Society of America*, vol. 90, pp. 1944–1957, 1991.

- [16] J. M. Weisenberger and M. E. Percy, "The transmission of phoneme-level information by multichannel tactile speech perception aids," *Ear and Hearing*, vol. 16, no. 4, pp. 392–406, 1995.
- [17] K. L. Galvin, G. Mavrias, A. Moore, R. S. Cowan, P. J. Blamey, and G. M. Clark, "A comparison of Tactaid II+ and Tactaid 7 use by adults with a profound hearing impairment," *Ear and Hearing*, vol. 20, no. 6, pp. 471–482, 1999.
- [18] E. Y. Wong, A. Israr, and M. K. O'Malley, "Discrimination of consonant articulation location by tactile stimulation of the forearm," in *2010 IEEE Haptics Symposium*, 2010, pp. 47–54.
- [19] R. Turcott, J. Chen, P. Castillo, B. Knott, W. Setiawan, F. Briggs, K. Klumb, F. Abnoui, P. Chakka, F. Lau, and A. Israr, "Efficient evaluation of coding strategies for transcutaneous language communication," in *Haptics: Science, Technology, and Applications*, D. Prattichizzo, H. Shinoda, H. Z. Tan, E. Ruffaldi, and A. Frisoli, Eds. Cham: Springer International Publishing, 2018, pp. 600–611.
- [20] M. Fontana de Vargas, A. Weill-Duflos, and J. R. Cooperstock, "Haptic speech communication using stimuli evocative of phoneme production," in *2019 IEEE World Haptics Conference (WHC) Tokyo, Japan*, 2019, pp. 610–615.
- [21] N. Dunkelberger, J. Sullivan, J. Bradley, N. P. Walling, I. Manickam, G. Dasarathy, A. Israr, F. W. Y. Lau, K. Klumb, B. Knott, F. Abnoui, R. Baraniuk, and M. K. O'Malley, "Conveying language through haptics: A multi-sensory approach," in *ISWC '18 Proceedings of the 2018 ACM International Symposium on Wearable Computers*, 2018, pp. 25–32.
- [22] S. Zhao, A. Israr, F. Lau, and F. Abnoui, "Coding tactile symbols for phonemic communication," in *CHI '18 Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, no. 392, 2018.
- [23] J. Chen, R. Turcott, P. Castillo, W. Setiawan, F. Lau, and A. Israr, "Learning to feel words: A comparison of learning approaches to acquire haptic words," in *Proceedings of ACM Symposium on Applied Perception (SAP'18)*, Vancouver, British Columbia Canada, 2018.
- [24] Y. Jiao, F. M. Severgnini, J. S. Martinez, J. Jung, H. Z. Tan, C. M. Reed, E. C. Wilson, F. Lau, A. Israr, R. Turcott, K. Klumb, and F. Abnoui, "A comparative study of phoneme- and word-based learning of English words presented to the skin," in *LNCS 10894*, D. Prattichizzo et al., Ed., 2018, pp. 623–635.
- [25] J. Jung, Y. Jiao, F. M. Severgnini, H. Z. Tan, C. M. Reed, A. Israr, F. Lau, and F. Abnoui, "Speech communication through the skin: design of learning protocols and initial findings," in *DUXU 2018*, LNCS 10919, A. Marcus and W. Wang, Eds., 2018, pp. 447–460.
- [26] H. Z. Tan, C. M. Reed, Y. Jiao, Z. D. Perez, E. C. Wilson, J. Jung, J. S. Martinez, and F. M. Severgnini, "Acquisition of 500 English words through a Tactile Phonemic Sleeve (TAPS)," *IEEE Transactions on Haptics*, 2020.
- [27] C. M. Reed, H. Z. Tan, Z. D. Perez, E. C. Wilson, F. M. Severgnini, J. Jung, J. S. Martinez, Y. Jiao, A. Israr, F. Lau, K. Klumb, R. Turcott, and F. Abnoui, "A phonemic-based tactile display for speech communication," *IEEE Transactions on Haptics*, vol. 12, no. 1, pp. 2–17, 2019.
- [28] J. C. Craig, "Pictorial and abstract cutaneous displays," in *Cutaneous Communication Systems and Devices*, F. Geldard, Ed. Austin, TX: Psychonomic Society, 1973, pp. 78–83.
- [29] J. M. Loomis, "Tactile letter recognition under different modes of stimulus presentation," *Perception & Psychophysics*, vol. 16, no. 2, pp. 401–408, 1974.
- [30] S. Saida, Y. Shimizu, and T. Wake, "Computer-controlled tvss and some characteristics of vibrotactile letter recognition," *Perceptual and Motor Skills*, vol. 55, no. 2, pp. 651–653, 1982.
- [31] Y. Yanagida, M. Kakita, R. W. Lindeman, Y. Kume, and N. Tetutani, "Vibrotactile letter reading using a low-resolution tactor array," in *12th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 2004. HAPTICS '04, 2004, pp. 400–406.
- [32] H. Kim, C. Seo, J. Lee, J. Ryu, S. bok Yu, and S. Lee, "Vibrotactile display for driving safety information," in *IEEE Intelligent Transportation Systems Conference*, 2006, pp. 573–577.
- [33] J. Wu, J. Zhang, J. Yan, W. Liu, and G. Song, "Design of a vibrotactile vest for contour perception," *International Journal of Advanced Robotic Systems*, vol. 9, no. 166, 2012.
- [34] G. Arnold and M. Auvray, "Perceptual learning: tactile letter recognition transfers across body surfaces," *Multisensory Research*, vol. 27, no. 1, pp. 71–90, 2014.
- [35] Y.-C. Liao, Y.-L. Chen, J.-Y. Lo, R.-H. Liang, L. Chan, and B.-Y. Chen, "Edgevib: effective alphanumeric character output using a wrist-worn tactile display," in *UIST '16 Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, 2016, pp. 595–601.
- [36] M. Janidarmian, A. Roshan Fekr, K. Radecka, and Z. Zilic, "Designing and evaluating a vibrotactile language for sensory substitution systems," in *Wireless Mobile Communication and Healthcare*, P. Perego, A. M. Rahmani, and N. TaheriNejad, Eds. Cham: Springer International Publishing, 2018, pp. 58–66.
- [37] M. Janidarmian, A. Roshan Fekr, K. Radecka, and Z. Zilic, "Wearable vibrotactile system as an assistive technology solution," *Mobile Networks and Applications*, 2019.
- [38] J. C. Cunha and P. Nohama, "A novel instrumentation to investigate the alternative tactile communication through mechanical stimulation using CO₂ jets," in *VI Latin American Congress on Biomedical Engineering CLAIB 2014, Paraná, Argentina 29, 30 & 31 October 2014, IFMBE Proceedings*, A. Braidot and A. Hadad, Eds., vol. 49. Springer, Cham, 2015, pp. 35–38.
- [39] J. O. Wobbrock, B. A. Myers, and J. A. Kembel, "EdgeWrite: A stylus-based text entry method designed for high accuracy and stability of motion," in *UIST '03: Proceedings of the 16th annual ACM symposium on User interface software and technology*, 2003, pp. 61–70.
- [40] P. Ranjbar, D. Stranneby, C. Akner-Koler, and E. Borg, "Identification of vibrotactile morse code on abdomen and wrist," *International Journal of Engineering Technology and Scientific Innovation*, vol. 1, no. 4, pp. 351–366, 2016.
- [41] C. Seim, S. Reynolds-Haertle, S. Srinivas, and T. Starner, "Tactile taps teach rhythmic text entry: Passive haptic learning of morse code," in *ISWC '16*, 2016, pp. 164–171.
- [42] C. Seim, R. Pontes, S. Kadiveti, Z. Adamjee, A. Cochran, T. Aveni, P. Presti, and T. Starner, "Towards haptic learning on a smart-watch," in *ISWC '18: Proceedings of the 2018 ACM International Symposium on Wearable Computers*, 2018, pp. 228–229.
- [43] E. Pescara, T. Polly, A. Schankin, and M. Beigl, "Reevaluating passive haptic learning of morse code," in *ISWC '19: Proceedings of the 23rd International Symposium on Wearable Computers*, 2019, pp. 186–194.
- [44] M. A. Plaisier, D. S. Vermeer, and A. M. L. Kappers, "Learning the vibrotactile morse code alphabet," *ACM Transactions on Applied Perception*, vol. 17, no. 3 (9), 2020.
- [45] S. Schätzle, B. Weber, and B. Leichtmann, "Hands-free reading Braille with a vibrotactile wristband," in *Proceedings of the IADIS International Conference Interfaces and Human Computer Interaction 2017*, 2017, pp. 35–45.
- [46] P. L. Brooks and B. J. Frost, "Evaluation of a tactile vocoder for word recognition," *Journal of the Acoustical Society of America*, vol. 74, no. 1, pp. 34–39, 1983.
- [47] P. L. Brooks, B. J. Frost, J. L. Mason, and K. Chung, "Acquisition of a 250-word vocabulary through a tactile vocoder," *Journal of the Acoustical Society of America*, vol. 77, no. 4, pp. 1576–1579, 1985.
- [48] M. P. Lynch, R. E. Eilers, D. K. Oller, and L. Lavoie, "Speech perception by congenitally deaf subjects using an electrocutaneous vocoder," *Journal of Rehabilitation Research and Development*, vol. 25, no. 3, pp. 41–50, 1988.
- [49] R. Velázquez, O. Bazán, C. Alonso, and C. Delgado-Mata, "Vibrating insoles for tactile communication with the feet," in *2011 15th International Conference on Advanced Robotics (ICAR)*, 2011, pp. 118–123.
- [50] R. Velázquez and E. Pissaloux, "On human performance in tactile language learning and tactile memory," in *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, 2014, pp. 96–101.
- [51] —, "Constructing tactile languages for situational awareness assistance of visually impaired people," in *Mobility of Visually Impaired People*, E. Pissaloux and R. Velázquez, Eds. Springer, Cham, 2018, pp. 597–616.
- [52] M. Walker and K. B. Reed, "Tactile morse code using locational stimulus identification," *IEEE Transactions on Haptics*, vol. 11, no. 1, pp. 151–155, 2018.
- [53] J. Wu, Z. Song, W. Wu, A. Song, and D. Constantinescu, "A vibrotactile system for image contour display," in *IEEE International Symposium on Virtual Reality Innovation 2011, 19–20 March, Singapore*, 2011, pp. 145–150.
- [54] R. Velázquez, O. Bazán, J. Varona, C. Delgado-Mata, and C. A. Gutiérrez, "Insights into the capabilities of tactile-foot perception," *International Journal of Advanced Robotic Systems*, vol. 9, no. 5, 2012.
- [55] F. A. Geldard and C. E. Sherrick, "Multiple cutaneous stimulation:

The discrimination of vibratory patterns," *Journal of the Acoustical Society of America*, vol. 37, no. 5, pp. 797–801, 1965.

- [56] E. E. Gottheil, C. R. W., and C. E. Sherrick, "The discrimination of vibratory patterns on a tactile matrix," *Bulletin of the Psychonomic Society*, vol. 11, pp. 21–24, 1978.

Astrid M.L. Kappers studied experimental physics at Utrecht University, the Netherlands. She received the PhD degree from Eindhoven University of Technology. From 1989 till September 2012, she was with the Department of Physics and Astronomy, Utrecht University. From 2008–2012, she was head of the Human Perception Group of the Helmholtz Institute. In September 2012, she moved with her whole group to the Department of Human Movement Sciences, Vrije Universiteit Amsterdam,

the Netherlands. October 2018, she moved to Eindhoven University of Technology, where she works in three groups: Dynamics and Control, and, Control System Technology of the Department of Mechanical Engineering, and Human Technology Interaction of the Department of Industrial Engineering & Innovation Sciences. She was promoted to full professor in 2005. Her research interests include haptic and visual perception. In 2003, she won the prestigious VICI grant. She is/was a member of the editorial boards of *Acta Psychologica* (2006-present) and *Current Psychology Letters* (2000-2011) and is an associate editor of the *IEEE Transactions on Haptics* (2007-2011 and 2017 till present).

Myrthe A. Plaisier received her Masters degree in experimental physics in 2006 and her PhD in 2010, both from Utrecht University. She performed her PhD research on haptic perception in the group of prof. Astrid Kappers. Her thesis received the thesis award from the Dutch Psychonomics Society. In 2011 she received a Rubicon grant from the Netherlands organisation for scientific research (NWO) which allowed her to continue her research at Bielefeld University in Germany. In 2013 she received a VENI grant

from NWO to investigate haptic perception of objects at the Department of Human Movement Sciences at Vrije Universiteit Amsterdam. In 2018 she joined Eindhoven University of Technology to work on the development of assistive technology for persons with deafblindness as part of the European project SUITCEYES.