Detection of incomplete, self-relevant auditory information presented to the unattended ear

Yoshizawa, Tatsuya
Human Information Systems Laboratory, Kanazawa Institute of Technology

Remijn, Gerard Bastiaan
Human Information Systems Laboratory, Kanazawa Institute of Technology | International Education Center, Faculty of Design, Kyushu University

Kitamura, Takumi
Human Information Systems Laboratory, Kanazawa Institute of Technology

https://hdl.handle.net/2324/1912787

出版情報：Acoustical Science and Technology. 33 (3), pp.147-153, 2012-06-04. 日本音響学会

Acoust. Sci. & Tech. 33, 3 (2012)

PAPER

Detection of incomplete, self-relevant auditory information presented to the unattended ear

Tatsuya Yoshizawa[1], Gerard Bastiaan Remijn[1,2,*] and Takumi Kitamura[1]

[1]*Human Information Systems Laboratory, Kanazawa Institute of Technology, 7–1 Ohgigaoka, Nonoichi, 921–8501 Japan*
[2]*International Education Center, Faculty of Design, Kyushu University, Minami-ku, Shiobaru 4–1–9, Fukuoka, 815–8545 Japan*

(*Received 18 March 2011, Accepted for publication 26 September 2011*)

Abstract: Dichotic listening studies have shown that information relevant to listeners, such as their own name, can be recognized even when presented to the unattended ear. Here, we used a dichotic listening paradigm to explore whether Japanese listeners could identify their name in the unattended ear even when sensory information was incomplete. The results showed that Japanese listeners with family names of 3, 4, or 5 morae — a speech unit equivalent to a syllable in English — recognized their name in about 20–60% of the trials even when the first or the last mora of the name was omitted. The data further showed a name-final effect under the 4- and 5-morae conditions: name recognition significantly decreased when the last mora of the listener's name was omitted as compared with the omission of the first mora. A possible explanation for these results is that self-relevant information, even when incomplete, automatically draws attention to the supposedly unattended ear and that the listener's recognition of the information is more robust when its end part is presented.

Keywords: Dichotic listening, Name recognition, Shadowing, Attention

PACS number: 43.66.Rq [doi:10.1250/ast.33.147]

1. INTRODUCTION

In the present study, we explored whether listeners could recognize their name in cases where a) either the first or last part of their name was omitted, and b) the incomplete name was presented to the listener's unattended ear. For this purpose, we used a dichotic listening task. In a dichotic listening task (Fig. 1), the listener receives a different auditory signal, usually a spoken message, in each ear. The listener's task is to ignore the message in one ear and to shadow, i.e., verbally repeat, the message received in the other ear. Rather than the shadowing itself, which is used to direct the listener's attention to only one ear, the experimenter is mainly interested in recall or recognition of the message presented to the unattended ear. Commonly known as the 'cocktail party problem,' the question is to what extent can listeners pick up auditory information they are not paying attention to?

The first dichotic shadowing tests by Cherry and Taylor [1] showed that, from the message presented to the unattended ear, listeners usually cannot pick up semantic information, even after numerous presentations. One exception to this was first reported by Moray [2]: while shadowing, listeners could detect their own name in the unattended ear. Moray used irrelevant messages of 5.5 minutes with the subject's own name (SON; following [3]) inserted once, either at 4 or 5 minutes of shadowing. Under these conditions, name recognition occurred in one-third of the cases. A similar percentage (34.6%) was obtained by Wood and Cowan [4] under similar stimulus conditions as adopted by Moray [2]. Later studies on the influence of the listener's working memory capacity on the name detection rate strongly suggested that recognition of the SON occurs because it automatically draws the listener's attention away from the shadowing task to the to-be-ignored channel — in spite of the instructions [5,6]. Because of the behavioral relevance of the SON, top-down processes may facilitate recognition. Further support for attention capturing by the SON or other self-relevant stimuli came from studies of the visual modality [7–10].

Dichotic listening studies mimic real-life situations to a small extent in that there are two sources of sound entering the auditory system at the same time, while we pay attention to only one source. In many real-life ('cocktail

*e-mail: remijn@design.kyushu-u.ac.jp

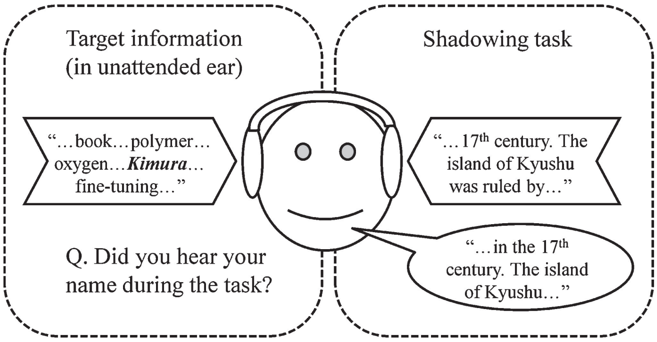

Fig. 1 The dichotic listening task used in the present study. A list of random words with the (in)complete target information (the listener's name) was presented to one ear, while at the same time, a story was presented to the other ear. The listener was required to verbally repeat ("shadow") the story word-by-word. After the task, the listener was asked whether he/she had heard the target information.

party') situations, though, our auditory system must deal with auditory sensory information that is strongly masked by the simultaneous presence of other sounds. Word recognition decreases with increasing levels of simultaneous masking (e.g. [11]). The robustness of word recognition has also been tested by completely erasing parts of a word. So-called "gating" studies have shown two effects. First, even when parts of a word are omitted, word recognition is more robust for highly familiar words than for unfamiliar words. Second, word recognition becomes more difficult if information at the beginning of the word is missing [12,13]. Namely, the first part (i.e., syllable) of a word triggers lexical access to a series of words stored in memory. As subsequent parts of the sensory signal enter the auditory system, the number of stored words with matching properties is narrowed down until a single word remains and is recognized. If the first part of the word is omitted, however, the triggering of lexical access is hampered.

Proceeding from the above findings, the question arises whether listeners can detect the SON in an unattended message even when the acoustic sensory information is incomplete, i.e., with missing syllables. How robust is recognition of the SON? Because of the behavioral relevance and familiarity of the SON, we hypothesized that even incomplete SON would trigger lexical access and thus facilitate name recognition. We further hypothesized that SON without the first part of sensory information would be more difficult to recognize than SON without the last part. That is, similar to the results of gating studies, the information at the beginning of the sensory signal would be more important for recognition than information at the end of the sensory signal.

Table 1 Target names of the participants who performed the experiment. Hyphens indicate mora transitions.

3-morae names:	4-morae names:	5-morae names:
i-zu-ka	o-o-ha-shi	wa-ka-ba-ya-shi
ko-sa-ka	a-ki-ya-ma	ka-ke-ba-ya-shi
ta-ka-se	na-ka-mu-ra	o-o-ka-wa-ra
ta-ma-ki	a-ka-sa-ka	
fu-ji-no	ka-wa-ka-mi	
o-o-no	ha-shi-mo-to	
su-zu-ki	ya-su-o-ka	

2. METHOD

2.1. Participants

Seventeen native-Japanese students, aged 20–24 years, voluntarily participated in the experiment. All had normal hearing and were naïve as to the purpose of the experiment. The names of the participants, the target stimuli in the experiment, are given in Table 1. The names of the listeners consisted of 3, 4, or 5 "morae." A mora is the building stone of the Japanese language, which is a syllable-timed language (e.g. [14]) spoken with an isochronous pattern. Since we intended to make the names incomplete, the use of Japanese names consisting of morae had a number of advantages. First, mora duration in spoken Japanese is relatively isochronous [15]. The average mora duration is 119 ms and the standard deviation ranges from 30–55 ms [16]. Targets consisting of (incomplete) Japanese family names thus would have a similar duration, proportional to the number of morae. Second, lexical and phrase accents are few with regard to Japanese names [17]. Even when cutting morae from the family names, the overall intonation and stress of the target information would thus remain relatively stable. By contrast, stress-accent languages like English vary considerably more in syllabic duration, with only 60% of the syllables falling between 106 and 260 ms [16].

As shown in Table 1, seven of the participants had a family name consisting of 3 morae and another seven had a family name consisting of 4 morae. Only three listeners with a family name consisting of 5 morae could be recruited, since such names are quite rare. For example, perusing a student attendance list consisting of 130 names resulted in only one name consisting of 5 morae. In comparison, 3-morae (57%) and 4-morae names (34%) were very common. A recent list of the 100 most common Japanese surnames published by The Japan Times also does not include 5-morae names [18].

2.2. Stimuli and Design

The stimuli were spoken messages of 30 s in duration. Shadowing sentences were presented to one ear, whereas

30 s random word lists — potentially containing a target — were presented to the other ear. The reason for using relatively short trials of 30 s was to increase the name detection rate. Under the conditions of full name presentation, this rate was about 33% with stimuli of 5.5 minutes [2,4,5]. In view of this, we expected that detection rates with incomplete name presentation would not exceed threshold values if we were to adopt similar stimulus conditions. Furthermore, contrary to previous studies, the same listeners were tested in multiple short trials.

There were 400 experimental stimuli. In half of these, the shadowing sentences were presented to the right ear and the random words to the left. In the other half, the ears of presentation were switched. The word lists could fulfill one of 4 target conditions. These targets were randomly inserted in the 30 s word list, but never in the first or final 5 seconds. Three of the conditions either comprised of the SON in full, the SON without the first mora, or the SON without the last mora (refer to Table 1). Under the fourth condition, the target in the word list was a randomly appointed word common in daily life and unrelated to the SON. There were 50 replications for each of the four target conditions. Sentences and words were arbitrarily taken from 50 different textbooks on five different topics: literature, economics, science, art, and history. In short, the 400 experimental stimuli were composed from 50 textbooks from which 4 target conditions were selected for both left-right or right-left presentations of the shadowing sentences and the random word list. Four-hundred control stimuli were also generated in the same manner as the 400 experimental stimuli. In the control stimuli, however, the random word list presented to the unattended ear did not contain (fragments of) the SON or any other appointed target word.

All stimuli were recorded by a native-Japanese male on a voice recorder (Olympus, Voice Trek V-13) and stored on and controlled by a personal computer (Dell Inspiron 580). The stimuli were recorded with a neutral voice and presented to the listener through headphones (Bose, QuietComfort3) in a sound-proof booth. The intensity level of the presented recordings was 67 A-weighted dB on average, as measured with a Rion Sound Level Meter (NL 32), against a background level of 34 dB in the booth.

2.3. Procedure

The 800 stimuli were randomly presented to the participant, who was asked to shadow the textbook sentences (i.e., verbally repeat the message) presented to the ear appointed by the experimenter. In half of the trials, the shadowing message was presented to the left ear, in the other half, to the right ear. In order to encourage the participant to focus solely on the shadowing task, he/she was told that their reading would be recorded and that he/she was therefore required to read in a clear voice. Only a single shadowing error was allowed. If the participant made more than one shadowing error, the trial was aborted and replaced with a new trial at the end of the session to complete the data set. Shadowing errors are an important indicator of switching attention between the ears [4], which could directly influence the results.

After each shadowing epoch, the participant was asked a number of questions about the word list presented to the unattended ear. One question was whether the participant had heard the SON. This question was interleaved with 7–10 dummy questions, in which the participant was asked whether he/she had heard certain objects, famous persons, or places being named in the unattended ear. The dummy questions were made in an effort to disguise the purpose of the experiment as much as possible and at least create uncertainty about the frequency of the targets with the (in)complete family name in the unattended ear. The experiment took 20 hours for each participant. Participants were paid for their time and allowed to schedule their participation in 1-hour sessions spanning 60 days. The procedure was approved by the Ethics Committee of the Kanazawa Institute of Technology and followed the declaration of Helsinki.

3. RESULTS

The number of times the listener detected the SON in the unattended message was expressed in proportions for each of the stimulus conditions. After arcsine transformation of the proportions, the data for the 3- and the 4-morae stimuli were subjected to two-way analyses of variance (ANOVA) with the stimulus condition (full SON, SON without the first mora, son without the last mora, or word unrelated to SON) and ear of presentation (left or right) as factors.

The two-way ANOVA for the 3-morae stimuli showed a significant main effect of stimulus condition [$F(3, 18) = 120.21$, $p < 0.01$]. Post-hoc Bonferroni tests ($p < 0.05$) showed that name recognition became significantly lower when the first or last mora was omitted from the SON as compared with full-name conditions (Fig. 2). Unsurprisingly, name recognition was significantly lower for the condition with a word unrelated to the SON as compared with presentation of the full SON, the SON without the first mora, and the SON without the last mora. Name recognition under the conditions without an actual SON was not exactly zero — some listeners even reported to have heard their name under those conditions. The ANOVA also showed a significant main effect of the ear of presentation [$F(1, 18) = 7.85$, $p < 0.05$]. Name recognition performance was better when the stimuli were presented to the right ear instead of the left. This has been observed repeatedly in dichotic listening studies [19,20].

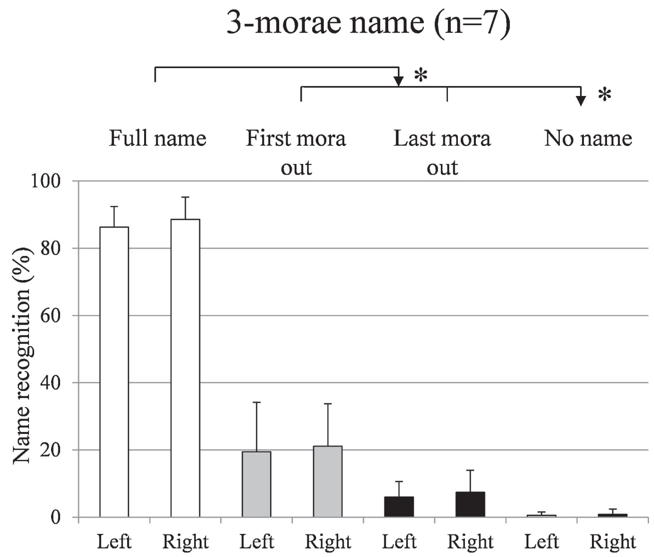

Fig. 2 Name-recognition percentage for the 3-morae stimuli. Error bars indicate the standard deviation.

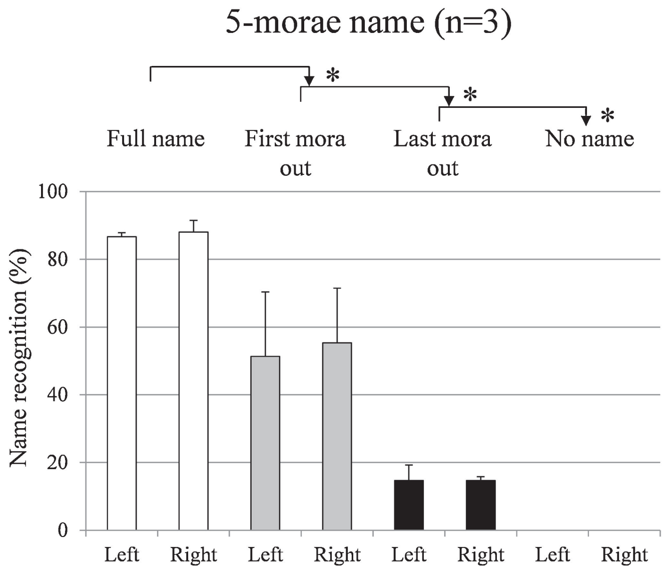

Fig. 4 Name-recognition percentage for the 5-morae stimuli. Error bars indicate the standard deviation.

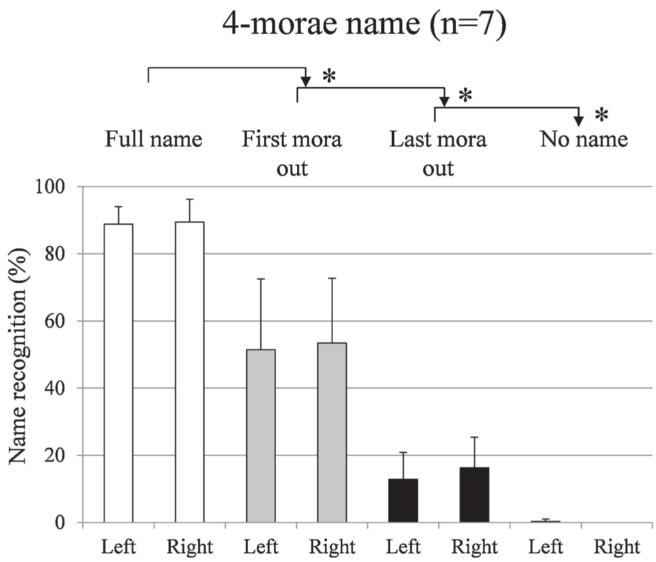

Fig. 3 Name-recognition percentage for the 4-morae stimuli. Error bars indicate the standard deviation.

The results for the 4-morae stimuli also showed a significant main effect of stimulus condition [$F(3, 18) = 173.69$, $p < 0.01$]. Bonferroni tests ($p < 0.05$) showed that name recognition was better for stimuli with (a part of) the SON than for the control stimuli unrelated to the SON. The highest name detection rate was obtained when the SON was presented, followed by stimuli in which the first mora was omitted from the SON (Fig. 3). Interestingly, the stimuli with the SON and the SON without the first mora yielded a significantly higher name detection rate than the stimuli without the last mora. The presentation of the last mora of the SON thus had more impact on name detection than the presentation of the first mora in this data set.

Because the 5-morae stimuli were presented to only three participants, we performed a one-way ANOVA over the stimulus condition. Data for ear of presentation were grouped together after a paired t-test over the collapsed data showed no particular left or right ear dominance ($t = 1.51$, $p < 0.14$). The results for the 5-morae stimuli showed the same trend as did those for the 4-morae stimuli. The main effect of the stimulus condition was significant [$F(3, 15) = 187.37$, $p < 0.01$], and post-hoc Bonferroni tests showed significant differences ($p < 0.05$) between all stimulus conditions (Fig. 4). Also in this data set, omission of the last mora of the SON impaired name recognition more than omission of the first mora.

4. DISCUSSION

The ANOVA-results are in line with those of earlier dichotic listening studies [2,4]; meaningful information, such as the SON, can be detected even when presented to the unattended ear. The results further confirm our first hypothesis in that Japanese listeners had the impression of hearing their name even when a part of it was omitted. Although name recognition was significantly worse for partial SON presentations than for full SON presentation, the rates were certainly not negligible. Proceeding from the literature suggesting that self-relevant information strongly activates the brain [21–23] and can draw attention to the to-be-ignored channel [4], the present data suggest that such an attention switch can even occur if the self-relevant information is incomplete.

A plausible explanation for the present results is that because of the behavioral relevance of the SON, it will not only attract the listener's attention when presented to the unattended ear, but it will also undergo top-down processing when incomplete. From classic masking studies, it is known that partial phonological information is more likely to be filled in if the absent phoneme is part of an existing word, rather than a nonexisting word [24]. This suggests that there is top-down feedback from partially activated lexical representations to earlier stages of lexical processing of auditory information [25]. The present results suggest that a part of the signal triggered stored knowledge about the SON, which was then fed back to facilitate recognition in the to-be-ignored ear.

The detection rate varied with the amount and the position of the omitted information from the SON. With regard to the relative amount of omitted information, the omission of one mora from a 3-morae SON resulted in a name detection rate of around 20% (Fig. 2). By contrast, when one mora was omitted from a SON of 4 or 5 morae (Figs. 3 and 4), name detection increased to 50–60%. Clues as to why this difference arose might be found outside the dichotic listening literature. First, listeners may have needed a certain absolute amount of acoustic information available to them for name detection to occur. Studies on word recognition in a foreign language have stressed the importance of absolute word length. For example, English listeners trying to detect target words in spoken French performed better with 4-syllable than with 1-syllable words [26]. Results of first-language word recognition have also shown an effect of word length [27,28]. If an effect of word length were transferable to the present study, it might be able to explain the increase in the name detection rate of incomplete, yet longer family names. Under conditions of one mora being left out, listeners with a 4- or 5-morae name would simply have had a longer absolute amount of information remaining to them than listeners with one mora left out of a 3-morae name.

An alternative explanation for the possible effect of name length on detection is that listeners, instead of an absolute duration, required a certain threshold percentage of the SON. Assuming that Japanese morae indeed all have similar durations [15], the difference between omitting 33% (3-morae), 25% (4-morae), or 20% (5-morae) of a SON, although relatively small, may have been of importance. Gating studies [29] have indeed suggested that a certain percentage of the acoustic information must be available to the listener in order for word recognition to occur. Listeners required 83% of a familiar word before they could identify it with certainty, — that is, under ideal listening conditions and regardless of word duration. The requirement of a threshold percentage of acoustic information necessary for detection of the SON, however, does not fit in with another major trend in the data. Namely, regardless of target size, the omission of the last mora had a more adverse effect on name detection than the omission of the first mora. Assuming that first and last morae have similar durations on average, in both cases, a similar percentage of information was left out.

The results thus did not confirm our second hypothesis, namely, that the omission of the first mora would impair name recognition more than the omission of the last mora. Why does the last mora have a relatively strong impact on name recognition? Studies on word recognition in English have shown that acoustic-phonetic information presented at the end of words is not crucial to the process of word recognition [30,31]. Rather, information at the beginnings of words could be more important [12,13]. Salasoo and Studdeft-Kennedy [32], for example, found that words gated from the beginning, thus increasing from the first phoneme, were recognized faster than words gated from the end. This advantage of forward gating over backward gating is in sharp contrast with the present results. Cultural factors cannot easily explain the name-final advantage on name recognition as found here. Japanese nicknames for first and family names, for example, typically consist of the first one or two mora(e) extended with a suffix that replaces the last mora(e) [33]. Name calling and recognition from childhood and later thus centers around the initial acoustic-phonetic information carried by the name. The name-final part is replaceable.

One far-fetched explanation for the relative importance of name-final information is related to attention switching between the ears and confirmation thereafter of whether the incoming information indeed matched with the SON. Under full name presentation, the first portion of the SON might cause the listener to switch attention to the to-be-ignored ear. Perceptual processing of the then attended sensory information takes time. Neuroimaging data have indicated that directing attention to one ear increases activation in structures modulating perceptual analyses of sound with a latency of about 150 ms [34]. The listener therefore attentively hears only the very last portion of the incoming information. If the last mora of the SON is missing, confirmation of whether the SON was actually presented might be difficult. By contrast, in stimuli in which the last mora is presented, the listener might match the information more confidently to his/her stored knowledge of the SON. However, in the case where the first mora is missing and the last mora is present, this explanation would only be valid if we assume that even the second mora of the SON can draw attention to the to-be-ignored ear and trigger its recognition.

Further research is necessary to clarify the name-final effect and to confirm the trend in the present data, since this could have interesting implications for theories of "early" versus "late" selection of auditory information processing in dichotic listening studies. If the detection of the SON in dichotic listening occurs because (a part of) the signal draws attention to the unattended ear [4–6], this could occur through "early" low-level stimulus properties in the signal, such as pitch or intensity changes or even a slight pause in the signal when the listener's (incomplete) name is inserted. Evidence supporting early selection had been described until recently (e.g. [35]). The present results, however, are not entirely compatible with the early-selection view. Omitting the first or the last mora should have had a rather similar effect on low-level properties. The results, though, showed a name-final effect. It is thus possible that name recognition might have occurred because higher-level linguistic properties of the SON [36] drew the listener's attention to the unattended ear and triggered recognition — this is a late-selection view. It must be noted, though, that because of the number of trials in the experiment our listeners could have become highly proficient in the dichotic listening task. They thus might even have divided their attention over both channels [6] to make use of (higher-level) linguistic properties for name recognition. This idea requires further research.

Other theoretical and procedural issues of the present study also need clarification. For example, we do not have much insight into how certain the listeners were that they had heard their name under various conditions. In order to validate the present results it therefore might be recommendable to use a signal detection approach that also concentrates on misses and false alarms. In order to apply such an approach, the design of the current study must be improved. Another suggestion for further research is to study the effect of the omission of the middle mora, where possible, to obtain a clearer view of the perceptual weight of the temporal position of a mora in the SON.

ACKNOWLEDGMENTS

We would like to explicitly state that all authors contributed equally to this study. The study was supported by MEXT grant 19530668 to TY and JSPS grant 22591278 to GR. We are very grateful to the two anonymous reviewers for their valuable insights and comments on this study.

REFERENCES

[1] E. C. Cherry and W. K. Taylor, "Some further experiments upon the recognition of speech, with one and with two ears," *J. Acoust. Soc. Am.*, **26**, 554–559 (1954).

[2] N. Moray, "Attention in dichotic listening: Affective cues and the influence of instructions," *Q. J. Exp. Psychol.*, **11**, 56–60 (1959).

[3] F. Perrin, P. Maquet, P. Peigneux, R. Perrine, C. Degueldre, E. Balteau, G. Del Fiore, G. Moonen, A. Luxen and S. Laureys, "Neural mechanisms involved in the detection of our first name: a combined ERPs and PET study," *Neuropsychologia*, **43**, 12–19 (2005).

[4] N. Wood and N. Cowan, "The cocktail party phenomenon revisited: How frequent are attention shifts to one's name in an irrelevant auditory channel?" *J. Exp. Psychol. Learn.*, **21**, 255–260 (1995).

[5] A. R. A. Conway, N. Cowan and M. F. Bunting, "The cocktail party phenomenon revisited: The importance of working memory capacity," *Psychon. Bull. Rev.*, **8**, 331–335 (2001).

[6] G. J. H. Colflesh and A. R. A. Conway, "Individual differences in working memory capacity and divided attention in dichotic listening," *Psychon. Bull. Rev.*, **14**, 699–703 (2007).

[7] D. J. Simons, "Attentional capture and inattentional blindness," *Trends Cognit. Sci.*, **4**, 147–155 (2000).

[8] D. E. Broadbent and M. H. P. Broadbent, "From detection to identification: Response to multiple targets in rapid serial visual presentation," *Percept. Psychophys.*, **42**, 105–113 (1987).

[9] K. L. Shapiro, J. Caldwell and R. E. Sorensen, "Personal names and the attentional blink: A visual "cocktail party" effect," *J. Exp. Psychol. Hum.*, **23**, 504–514 (1997).

[10] C. Frings, "Relevant distractors do not cause negative priming," *Psychon. Bull. Rev.*, **13**, 322–327 (2006).

[11] E. L. Goshorn and E. K. Robertson, "Effects of multiple background talkers on word recognition and response awareness," *J. Acoust. Soc. Am.*, **113**, 2289 (2003).

[12] W. D. Marslen-Wilsen and A. Welsh, "Processing interactions and lexical access during word recognition in continuous speech," *Cognit. Psychol.*, **10**, 29–63 (1978).

[13] A. Salasoo and D. B. Pisoni, "Interaction of knowledge sources in spoken word identification," *J. Mem. Lang.*, **24**, 210–231 (1985).

[14] M. Beckman, "Segment duration and the "mora" in Japanese," *Phonetica*, **39**, 113–135 (1982).

[15] N. Warner and T. Arai, "Japanese mora-timing: A review," *Phonetica*, **58**, 1–25 (2001).

[16] T. Arai and S. Greenberg, "The temporal properties of spoken Japanese are similar to those of English," *Proc. Eurospeech*, Rhodes, Greece, Vol. 2, pp. 1011–1014 (1997).

[17] H. Sato, "Interaction between phonetic features and accent-placement in Japanese family names," in *Prosody and Syntax: Cross-linguistic Perspectives*, Y. Kawaguchi, I. Fónagy and T. Moriguchi, Eds. (John Benjamins Publishing, Philadelphia, 2006), pp. 223–238.

[18] Japan's top 100 most common family names (2009, Oct. 11). Retrieved Oct. 20, 2009 from http://www.japantimes.co.jp.

[19] D. Kimura, "A note on cerebral dominance in hearing," *Acta Otolaryngol.*, **56**, 617 (1963).

[20] R. H. Wilson and M. S. Jaffe, "Interactions of age, ear, and stimulus complexity on dichotic digit recognition," *J. Am. Acad. Audiol.*, **7**, 358–364 (1996).

[21] I. Holeckova, C. Fischer, D. Morlet, C. Delpuech, N. Costes and F. Maugière, "Subject's own name as a novel in a MMN design: A combined ERP and PET study," *Brain Res.*, **1189**, 152–165 (2007).

[22] K. K. Kampe, C. D. Frith and U. Frith "Hey John: Signals conveying communicative intention toward the self activate brain regions associated with "mentalizing," regardless of modality," *J. Neurosci.*, **23**, 5258–5263 (2003).

[23] M. Miyakoshi, M. Nomura and H. Ohira, "An ERP study on self-relevant object recognition," *Brain Cognit.*, **63**, 182–189 (2007).

[24] R. M. Warren, "Perceptual restoration of missing speech sounds," *Science*, **167**, 392–393 (1970).

[25] E. M. Saffran and M. F. Schwartz, "Language," in *Handbook of Psychology*, M. Gallagher, R. J. Nelson and I. B. Weiner, Eds. (John Wiley & Sons, Hoboken, N.J., 2003), pp. 609.

[26] H. Goldstein, "Word recognition in a foreign language: a study of speech perception," *J. Psycholinguist. Res.*, **12**, 417–427 (1983).

[27] C. Craig and B. Kim, "Effects of time gating and word length

on isolated word-recognition performance," *J. Speech Hear. Res.*, **33**, 808–815 (1990).

[28] M. A. Pitt and A. G. Samuel, "Word length and lexical activation: Longer is better," *J. Exp. Psychol. Hum.*, **32**, 120–135 (2006).

[29] F. Grosjean, "Spoken word recognition processes and the gating paradigm," *Percep. Psychophys.*, **28**, 267–283 (1980).

[30] W. D. Marslen-Wilson, "Linguistic structure and speech shadowing at very short latencies," *Nature*, **244**, 522–523 (1973).

[31] S. Cotton and F. Grosjean, "The gating paradigm: A comparison of successive and individual presentation formats," *Percept. Psychophys.*, **35**, 41–48 (1984).

[32] A. Salasoo and M. Studdeft-Kennedy, "Some processes of word recognition in fluent speech," *J. Acoust. Soc. Am.*, **71**, 95 (1982).

[33] S. Tanno and K. Maeda, "University students and nicknames: The origins of nicknames and the feelings toward nicknames," *Hiroshima Psychol. Res.*, **7**, 311–314 (2007) (in Japanese with English abstract).

[34] B. Ross, S. A. Hillyard and T. W. Picton, "Temporal dynamics of selective attention during dichotic listening," *Cereb. Cortex*, **20**, 1360–1371 (2009).

[35] E. Dupoux, S. Kouider and J. Mehler, "Lexical access without attention? Explorations using dichotic priming," *J. Exp. Psychol. Hum.*, **29**, 172–184 (2003).

[36] J. A. Deutsch and D. Deutsch, "Attention: Some theoretical considerations," *Psychol. Rev.*, **70**, 80–90 (1963).