Phonological processing differences in bilinguals and monolinguals

Henna Tamminen a,b,c,*, Maija S. Peltola a,b,c, Heidi Toivonen a,b, Teija Kujala d,e, Risto Näätänen e,f,g

a Department of Phonetics, University of Turku, Finland
b Centre for Cognitive Neuroscience, University of Turku, Finland
c LAB-lab, Department of Phonetics, University of Turku, Finland
d Cicero Learning Network, University of Helsinki, Finland
e Cognitive Brain Research Unit, Institute of Behavioural Sciences, University of Helsinki, Finland
f Center of Functionally Integrative Neuroscience, University of Aarhus, Denmark
g Department of Psychology, University of Tartu, Estonia

1. Introduction

A variety of linguistic backgrounds may result in bilingualism and the definition may be based on the age and timing of the acquisition of the two languages (from birth simultaneously or later in life consecutively) or the obtained level of competence (Bloomfield, 1962; Grosjean, 1989). This undoubtedly leads to differing, even contradictory, results in studies on bilingual speech perception. For example, it has been shown that the two languages of early, fluent bilinguals are neurally represented in the same brain areas (Chee et al., 1999). In contrast, Klein et al. (1995) showed that this applies also when the languages are learned later in life. Perani et al. (1996) found that languages are represented in separate brain areas in less fluent later bilinguals, whereas Perani et al. (1998) found that the neural activations are similarly distributed in fluent bilinguals (early or late). They suggested that the localisation of cortical representations of the second language is determined more by the attained proficiency than the age of acquisition. A more complex pattern was found by Kim et al. (1997) who showed that, in the frontal lobe, early learners' languages are represented in a shared area while late learners' languages are spatially separated; however, there were no such distinctions in the activations in the temporal lobe. Different learning backgrounds also lead to differences in speech sound representations: new memory traces for non-native speech sounds evolve in cases of immigration (Winkler et al., 1999), but not in classroom (Peltola et al., 2003). Furthermore, these learning contexts result in either separate or intertwined phonological systems: Immigrants (Winkler et al., 1999) seem to have functionally inseparable languages while classroom learning (Peltola & Aaltonen, 2005; Peltola et al., 2012) leads to the development of functionally distinct phonemic systems. These kinds of differences in the results may partly be explained by the use of different methods, varying stimulus selection criteria, and different definitions of bilingualism. The distinction between balanced and dominant bilinguals proposed by Albert and Obler (1978) offers a solution to the problematic definition of bilingualism. Learning background may be one criterion when classifying bilinguals into balanced and dominant ones – balanced having acquired both languages from birth, in a one-language–one-parent setting, and dominants having learned the second language later in life. This dichotomy presupposes that the age of acquisition and proficiency level are natural features of the type of bilingualism.

Speech sound perception is known to be language-specific and mother tongue phonemes are discriminated easily and preattentively on the basis of long term memory traces (Näätänen et al., 1997), and these memory traces and the consequent phonological system develop early in childhood (Cheour et al., 1998). Behavioural (Kuhl, 1991; Liberman et al., 1957) and psychophysiological (Sharma & Dorman, 1989) studies have shown that discrimination sensitivity is highest in bilinguals who acquired both languages from birth.
at the native language phoneme boundaries and weakest near the native speech sound category prototypes. In behavioural discrimination this is reflected, e.g., in shorter vs. longer reaction times and in preattentive neural processing native-likeness is connected with responses having shorter latencies and larger amplitudes (Kujala & Näätänen, 2010). This is problematic in bilingual speech perception: balanced bilinguals’ discrimination sensitivity peaks at the category boundary of one language while this boundary may be near a prototype of the other language which impedes discrimination. For example, within the closed vowel continuum Finnish distinguishes between three categories (/i/–/y/–/u/) while Swedish has four vowels (/i/–/y/–/u/–/u/). Therefore, the category boundary between the Finnish /y/ and /u/ is located within the Swedish category /u/. For balanced Finnish–Swedish bilinguals this means that the same acoustic distinction may be easy or difficult to discriminate depending on whether they rely on the Finnish or Swedish phonology.

The goal of the present study was to determine whether Balanced Finnish–Swedish bilingual speech processing is different from Monolingual Finnish perception. We aim to find out if the co-existence of two phonological systems results in different kinds of perceptual processings in bilinguals, compared to monolinguals who have only one native language, and how this potential difference is manifested in preattentive memory trace retrieval. Memory trace activations can most conveniently be studied by using the mismatch negativity (MMN) component of the event-related potential (ERP) since the MMN allows us to measure how fast (latency) and strongly (amplitude) memory traces are accessed, strong memory traces being reflected in short-latency and large-amplitude MMNs (Kujala & Näätänen, 2010). Our hypothesis is that speech processing may be different in bilinguals in comparison with monolinguals who have only one native language, and how this potential difference is manifested in preattentive memory trace retrieval.

Two groups of voluntary, right handed (tested with Edinburgh Handedness Inventory (Oldfield, 1971)), normally hearing (tested with an audiometer with perceptually relevant frequencies 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz), and neurologically healthy subjects participated in this study. The first group consisted of 10 (mean age 26.7 years, 7 females) native speakers of Finnish (Monolinguals) and the second group consisted of 12 (mean age 20.3 years, 7 females) Finnish–Swedish bilinguals (Balanced bilinguals). The Balanced bilinguals were from the same Finland–Swedish dialectal area, they had acquired Finnish and Finland–Swedish from birth (one-language–one-parent), and none of them had ever lived in Sweden. They reported a high proficiency in both languages and they continued using both languages equally often in their daily lives. The socioeconomic status of these languages in Finland is equal as they are official languages and all public services are provided with both languages which enable concurrent use of both. The bilingual subjects and their data are the same as in Peltola et al. (2012) study. Prior to testing, a written consent was obtained from the subjects, and the experiments were conducted according to the guidelines defined by the ethical committee of the University of Turku.

The behavioural identification test (forced choice) consisted of 18 synthetic (HLsyn software, version 1.0. Sensimetrics, Inc.) isolated vowels from the closed rounded vowel continuum which is divided into two categories in Finnish, /y/–/u/, and three in Swedish, /i/–/u/–/u/. The values for the second formant (F2) ranged from 606 Hz (703 Mel) to 2077 Hz (1553 Mel) in steps of 50 Mel while the other formants were kept constant (F1 = 250 Hz, F3 = 2600 Hz, and F4 = 3500 Hz). Fig. 1 shows an illustration of the continuum and roughly the category boundaries in question. It should be noticed that, the Finland–Swedish and Sweden–Swedish closed vowels are somewhat different, especially the close central /u/ has a lower F2 in Finland–Swedish (which is the one in question in this study) than in Sweden–Swedish (Asu et al., 2009). To imitate natural speech, the fundamental frequency (F0) contour was set at 112 Hz at the onset of the stimuli then reaching the maximum 132 Hz by 100 ms and finally descending to 92 Hz at the end. The stimuli contained a 30 ms ramp both at the onset and at the offset during which the amplitude was smoothed. The duration of the stimuli was 350 ms.

In the EEG recordings we used an individually selected stimulus pair for each subject (the stimuli always being 100 Mels apart) to ensure that the stimulus contrast had a phonological status in one of the participant’s languages (Finnish), but not in the other (Swedish). This may not have necessarily been the case, if we had used group average stimuli, since the acoustic area of the closed vowels is quite vast. The stimulus pair was chosen on the basis of identification experiments carried out prior to the MMN registrations. In these ID-tests we located the /y/–/u/ category boundary for Finnish (from both groups) as well as the /y/–/u/ and /u/–/u/ boundaries for Swedish (from Balanced bilinguals) individually for each subject and ensured that the Finnish boundary was located within the Swedish category /u/.

Fig. 1. The closed rounded vowels used in the experiment. The extremes are at 2077 Hz (/y/) and 606 Hz (/u/), and the mid-vowel is roughly placed on the most /-like location. The category boundaries are also indicated by arrows.
mean Finnish /y/–/u/ category boundary for the Monolinguals was between stimuli 9 (F2: 1163 Hz) and 10 (F2: 1247 Hz) at point 9.7 (standard deviation 0.8) in the 18 vowel continuum. For the Balanced bilinguals the same boundary was near stimulus 9 (F2: 1163 Hz) at 8.8 (standard deviation 2.5). However, at the same continuum, the Balanced bilinguals located the Swedish /y/–/u/ category boundary at 12.9 (F2 for stimulus 13 was 1524 Hz; standard deviation 1.2) and /y/–/u/ boundary between stimuli 4 (F2: 792 Hz) and 5 (F2: 860 Hz) at 4.6 (standard deviation 0.7). Obviously, monolingual Finns received all instructions for the ID experiment and MMN registration in Finnish. Bilinguals, in turn, participated twice in the identification experiment so that, within one session, only Finnish was used and they were instructed to label the vowels as Finnish /y/ and /u/ while, within the other, the instructions (label the vowels as Swedish /y/, /u/ and /u/) were provided in Swedish by a native speaker. During the completely monolingual Finnish session, we also registered the MMN responses to the Finnish contrast. The spectrograms of the most often used stimulus pair for the Balanced bilinguals are illustrated in Fig. 2, and further, the minimum, maximum and most often used F2 values are shown in Table 1.

We registered nose-referenced electroencephalogram (EEG) (Synamps amplifier; sampling rate 250 Hz, bandwidth 0.5–70 Hz) in the oddball paradigm. The stimuli were presented pseudorandomly, with a representative of the /y/ category as the standard and an exemplar of /u/ as the deviant (783 standards, 120 deviants, deviant probability 0.13). The inter-stimulus interval was 550 ms. The EEG was recorded from the scalp with Sn electrodes (a 21 channel electrocap, Electro-Cap International, Inc.) and eye movements were monitored with two electrooculogram (EOG) electrodes attached below and near the outer canthus of the right eye. Left mastoid impedance was kept under 5 kΩ.

The event-related potential (ERP) epochs (a 600 ms window in- cluding a 100 ms pre-stimulus baseline period) were digitally filtered off-line with a 1–30 Hz bandpass filter and the artefact criterion was set at ±100 μV. Separately averaged waveforms for the standard and the deviant stimuli were computed for each subject and difference waveforms were created by subtracting the response to the standard stimulus from that to the deviant stimulus. We selected two consecutive 50 ms time windows (180–230 ms and 230–280 ms) around the maximum amplitudes in the difference waveforms. Six electrodes (Fz, Cz, F4, C4, C3, C4) were selected for the statistical analysis of the mean amplitudes. The MMN peak latency was automatically measured from the Fz electrode at 150–300 ms from stimulus onset, however, it was also manually ensured that there was a negative peak in that time frame for each subject. MMN latency differences can reliably be studied at Fz electrode because generally the component is most clearly obtained at that site (e.g. Kujala et al., 2007). The mean amplitudes were statistically analysed using a Group (2)×Time window (2)×Electrode (6) Repeated Measures analysis of variance (ANOVA). The required further post-hoc tests were separately carried out for each time window (Group (2)×Electrode (6) ANOVA). The latency data were analysed with a Group (2)×latency one-way analysis of variance (ANOVA).

3. Results

Our results showed that the groups differed in the MMN latency (Fig. 3). The mean peak latencies analysed from a 150–300 ms time window from Fz were 218.4 ms (standard deviation 38.7) for the Monolingual group and 250.67 ms (standard deviation 30.8) for the Balanced bilingual group. The latency difference can be seen in two different analyses. First, the peak latency analysis proper showed a significant main effect of group (F(1,20) = 4.754, p = 0.041). Second- ly, the latency difference is indirectly shown in the mean amplitude analysis where there was also a significant interaction between group and selected time window (180–230 ms vs. 230–280 ms) (F(1,20) = 8.399, p = 0.009). The further post-hoc tests showed that the amplitude difference was valid in the first time window (F(1,20) = 12.177, p = 0.002), where the Monolinguals’ change detection response was larger, whereas no group differences were found in the second time window where both groups had an extensive response.

Furthermore, the MMN amplitudes also differed between the groups as the mean amplitude analysis revealed the main effect of group (F(1,20) = 4.834, p = 0.040) suggesting an overall reduced MMN amplitude in the Balanced bilingual subjects. The fact that this group difference was valid only in time window one, but not in time window two, suggests not only the difference in latency but also in amplitude, since the overall mass of the Monolingual response seems more extensive (see Fig. 3 for visual support). The mean MMN amplitudes for both time windows in six electrodes are displayed in Table 2.

Please cite this article as: Tamminen, H., et al., Phonological processing differences in bilinguals and monolinguals, Int. J. Psychophysiol. (2012), http://dx.doi.org/10.1016/j.ijpsycho.2012.10.003

![Fig. 2. Spectrograms of the most often used stimulus pair for the Balanced bilinguals. The representative of the /y/ category is on the left and /u/ on the right. The F2 value for /y/ is 1163 Hz and for /u/ 1005 Hz.](image-url)
4. Discussion

The present study aimed at determining whether balanced bilingual speech processing differs from monolingual speech processing. It could be speculated that, if Balanced bilinguals could switch off one of their two languages when the situation required the use of only one language, their MMN response may be identical with the Monolinguals’ response, both in amplitude and latency. If, on the other hand, both their native languages are active all the time, the non-target language could affect the processing of the target language.
language. We tested this by comparing the MMNs for a Finnish vowel contrast in Finnish–Swedish Balanced bilinguals and Monolingual Finns.

Our results suggest that, while early bilinguals and monolinguals do not differ in their processing of non-linguistic stimuli (Nenonen et al., 2003; Ortiz-Martíllia et al., 2010), the preattentive processing of speech sounds in Balanced bilinguals is different from Monolinguals. The main finding was that the MMN latency was longer, and it also seemed that the overall amplitude was smaller, in the Balanced bilinguals than in the Monolingual subjects. In general, the MMN latency decreases as the deviation of the stimuli increases (e.g. Näätänen et al., 2005) and, on the other hand, also training of non-native contrasts decreases latency (Menning et al., 2002).

Surprisingly, unlike in training, it seems that when there are two native languages, like in balanced bilinguals, compared to one native language, the MMN latency increases as there are more phonological categories to process. This may be caused by the more extensive amount of potentially retrievable categories. These results suggest that, compared to Monolinguals, Balanced bilinguals may have a more extensive intertwined phonological system where both languages are active all the time. As a consequence, individual phonological items and their memory traces are accessed more slowly (see also Peltola et al., 2012). This slowing effect has been reported in lexical learning by Davis et al. (2008) in a study where they showed that lexical competition between similar-sounding existing words and newly acquired words resulted in slower lexical access. In other words, a more extensive vocabulary inventory slowed down memory trace retrieval. This is in line with our finding of the longer-latency response in bilinguals, which, therefore, suggests that the systems are intertwined with both phonologies available at the same time. On the other hand, the tentative result showing a reduced MMN amplitude, i.e. a non-native-like response, in the bilinguals could be explained by interference from the other language, where the used contrast is non-phonemic. This would again mean that both systems are active during the processing of only one of the languages. Taken together, the slower and smaller MMN in Balanced bilinguals may be linked with the competition between these two native-like phonological systems. The phonologically relevant Finnish /æ/ – /œ/ vowel pair is located within the Swedish vowel category /ai/, i.e. the same acoustic contrast is relevant in one language and irrelevant in the other. This dual role may explain the prolonged latency and the non-native-like response in the bilingual subjects.

Acknowledgements

The study was financially supported by Emil Aaltonen Foundation and the Academy of Finland (nos. 206352, 128840, and 122745). The authors thank all the subjects who participated, Jyrki Tuomainen, PhD for his help at the beginning of this project, and Heli Kurttila and Jaana Hirvelä, M.A. for their assistance in data collection.

References


Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inven-
yory. Neuropsychologia 9, 97–113.


Sharma, A., Dorman, M.F., 1999. Cortical auditory evoked potential correlates of categori-


Please cite this article as: Tamminen, H., et al., Phonological processing differences in bilinguals and monolinguals, Int. J. Psychophysiol. (2012), http://dx.doi.org/10.1016/j.ijpsycho.2012.10.003