PSYCHOLOGY & NEUROSCIENCE

Psychology & Neuroscience, 2014, 7, 2, 183 - 191
DOI: 10.3922/j.psns.2014.015

Electrophysiological investigation of the functional overlap between semantic and equivalence relations

Renato Bortoloti[1,2], Naiene Pimentel[1,2], and Julio C. de Rose[1,2]

1. Universidade Federal de São Carlos, São Carlos, SP, Brazil

2. Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino, São Carlos, SP, Brazil

Abstract

Recent research using the event-related potential (ERP) technique has shown that equivalence relations have properties similar to genuine semantic relations. This study aimed to advance electrophysiological investigations of the functional overlap between semantic and equivalence relations. The N400 component, an index of semantic processing, was used to measure whether semantic relations were experimentally established between arbitrary stimuli. The stimuli became equivalent via a matching-to-sample training designed to maximize the establishment of equivalence relations and the strength of the classes. Non-equivalent pairs of stimuli elicited larger N400 responses than equivalent pairs in electrodes placed over the central and parietal scalp regions, providing additional support for the assumption that stimulus equivalence is an appropriate model of semantic relations. Latency of the N400 component was shorter than in previous studies, probably due to experimental parameters that maximized relational strength. These data raise the possibility that N400 latency may provide a continuous measure of relational strength, thus supplementing the all-or-none character of equivalence tests based on matching to sample.
Keywords: stimulus equivalence, event-related potentials, semantic relationship.

Received 14 February 2014; received in revised form 14 March 2014; accepted 02 April 2014. Available online 27 June 2014.

When humans match dissimilar stimuli (e.g., words to pictures), equivalence relations between those stimuli are likely to be demonstrated through behavioral tests derived from the logical properties of reflexivity, symmetry and transitivity. If these properties are confirmed, one can say that those stimuli constitute a class of equivalent stimuli in which each member is substitutable for the others (Sidman & Tailby, 1992). Sidman (1994) argued that equivalence relations are important phenomena in their own right, and also that they help to explain properties of human language. The mutual substitutability implied in the equivalence paradigm could be used to demonstrate, for instance, "one way that symbols do become established as such, one way that words can come to 'mean' what they 'stand for'" (Sidman, 1994, p. 563).

Earlier investigations on transfer of functions (e.g., Barnes-Holmes, Keane, Barnes-Holmes, & Smeets, 2000; de Rose, McIlvane, Dube, Galpin, & Stoddard, 1988, Dougher, Augustson, Markham, Greenway, & Wulfert, 1994; Hayes, Kohlenberg, & Hayes, 1991), contextual control (Bush, Sidman, & de Rose, 1989), class fusion (Sidman & Tailby, 1982), and class expansion (Saunders, Saunders, Kirby, & Spradlin, 1988) tended to confirm the assertion that the stimulus equivalence paradigm allows a behavioral approach of symbolic processes. More recently, the validity of stimulus equivalence as a behavioral model of symbol and meaning has been supported by studies involving different techniques such as semantic differential (Bortoloti & de Rose, 2009), semantic priming (Barnes-Holmes et al., 2005; Bortoloti & de Rose, 2011a), Implicit Association Test (O'Toole, Barnes-Holmes, & Smyth, 2007), and Implicit Relational Assessment Procedure (Bortoloti & de Rose, 2012).

Recent research using the event-related potential (ERP) technique has shown that equivalence relations have properties similar to genuine semantic relations (Barnes-Holmes et al., 2005; Haimson, Wilkinson, Rosenquist, Ouimet, & McIlvane, 2009). ERPs are small changes in brain electrical activity evoked by some external or internal event and recorded from the scalp. By the fact that the ERP is superimposed on the electroencephalogram (EEG) in progress, it is often difficult to see an ERP from a single presentation of a particular event. Typically, a reasonable number of similar events is repeated and EEG epochs time-locked

Renato Bortoloti, Naiene Pimentel, and Julio C. de Rose, Universidade Federal de São Carlos and Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino. Correspondence regarding this article should be directed to: Julio C. de Rose, Universidade Federal de São Carlos, Rodovia Washington Luís Km 235, SP-310, 13570000 - São Paulo, SP, Brasil. Phone: +55 (16) 3351 8492. E-mail: julioderose@gmail.com

to these events are averaged. Regular patterns of ERPs, called *components*, may be associated with specific functional processes. Components are discriminated between positive and negative fluctuations, which are often designated by their polarity and latency or ordinal position after stimulus presentation. For example, the N400 is a negative deflection (N for negative) that peaks about 400 ms after the presentation of a stimulus. The P3 is the third positive peak (P for positive) that appears after the presentation of a stimulus. A number of language-related components has been characterized as regularly evoked by, for instance, acoustic-phonetic, phonological, orthographic, semantic, syntactic, and prosodic events (Steinhauer & Connolly, 2008).

Linguistic research involving ERPs has characterized the "semantic component" N400 (Kutas & Hillyard, 1980). The N400 is a voltage fluctuation that is prominent over central and parietal regions of the scalp. The component reflects incompatibility between a stimulus and a previously established semantic context. For instance, a large N400 would be evoked by the last word of the sentence "I had a coffee with sugar and *nose*", but not by the last word in the sentence "I had a coffee with sugar and *cream*". Similar results are obtained for pairs of words presented in sequence, in *semantic priming* procedures, in which a large N400 is elicited by the second word if it is incompatible with the semantic context established by the first word (e.g., a large N400 is evoked by the second word in *coffee-nose*, but not in *coffee-cream*)[1].

Barnes-Holmes et al. (2005) and Haimson et al. (2009) showed that the N400 component can be useful in the study of neural correlates of the formation of equivalence classes. A brief description of these studies will help to contextualize our own goals.

Barnes-Holmes et al. (2005, Experiment 3) trained undergraduate students to establish two 4-member equivalence classes through a linear training protocol AB, BC, and CD. The stimuli were word-like nonsense words. Then, participants were exposed to a "lexical decision task" in which two nonsense words (prime and target stimuli) remained on screen together. The authors found that stimulus pairs that did not belong to the same equivalence class evoked a larger N400 component compared to pairs that belonged to the same equivalence class. Also searching for an electrophysiological signature for stimulus equivalence, Haimson et al. (2009, Experiment 1) first replicated the N400 component with pairs of common English words (i.e., non-related common words evoked a larger N400 than related ones). Then, Haimson et al. (2009, Experiment 2) taught other participants to establish relations AB, AC, AD, AE, and AF comprising arbitrary forms and exposed these participants to a "semantic priming procedure" involving the stimuli B, C, D, E and F. It was found that the presentation of stimulus pairs such as BC, DE, FD, etc., in which the first stimulus simulated a semantic context for the second one, evoked a large N400 component only when these stimuli did not belong to the same equivalence class (e.g., pairs such as B1C2, D2E3, F3D1, etc.). Findings from both studies gave, thus, a new type of support for stimulus equivalence as a model of semantic meaning: it was showed that the N400 component is obtained both in experimental conditions that involve semantic relations of the participants' native language and in experimental conditions involving arbitrary equivalence relations.

The present study aimed to advance the electrophysiological investigation of the functional overlap between semantic and equivalence relations. We also used the N400 technique to index relations between arbitrary visual stimuli that became equivalent via a matching-to-sample training procedure. The procedure adopted in the current study involved experimental parameters that differ from those adopted by the former ones in some important ways. First, the stimulus presentation during the electrophysiological recording differs from the stimulus presentation employed by Barnes-Holmes et al. (2005, Experiment 3). Barnes-Holmes et al. used a variant of a lexical decision task in which prime and target stimuli remained on screen together with the prime located above the target. The prime was presented 100 ms before the target, but both remained together for 1,750 ms. In the current study, we selected a single-stimulus paradigm in which only one stimulus was presented at a time. We believe that a single-stimulus paradigm can improve our ability to time-lock the ERPs to the instant that the target stimulus was shown, i.e., the moment in which the perceptual and cognitive processes in which we are interested are evoked. We also consider that a single-stimulus paradigm can help the establishment of a more stable baseline prior the presentation of the target stimulus.

Second, relational training was designed to maximize the establishment of equivalence relations and the strength of the classes established. We used a sample-as-node training protocol, which is considered more efficient in generating equivalence classes than the linear training protocol (e.g., Arntzen & Holth, 1997). We also used a matching to sample training with a 2-s delay between the removal of the sample and the presentation of the comparison stimuli. The delayed matching to sample is a parameter that enhances the formation of equivalence classes, as shown by Arntzen (2006), Vaidya and Smith (2006), and Bortoloti and de Rose (2009, Experiment 2). Besides influencing the formation of equivalence classes, delayed matching to sample can also strengthen the classes established, determining

[1] The N400 has been considered as a useful indicator of semantic relations in a number of experimental circumstances (e.g., Kutas & Federmeier, 2000), but this component probably does not index a monolithic process. Some ERP experiments have found N400 in tasks not related to language events (e.g., Niedeggen & Rösler, 1996), and even in tasks clearly independent of the lexical-semantic domain (e.g., Bornkessel, Schlesewsky, & Friederici, 2002; Osterhout, 1997).

higher transfer of functions among equivalent stimuli (Bortoloti & de Rose, 2009; 2011b; 2012).

Third, participants were taught to establish two 5-member equivalence classes including both meaningful and arbitrary stimuli. The meaningful stimuli were pictures of human faces expressing either anger or happiness. During the sample-as-node-matching-to-sample training, faces were always the samples (A) and arbitrary forms were always the comparisons (B, C, D, and E). The samples were members of perceptual classes (Fields et al., 2002; Fields & Moss, 2008): The stimuli designated as A1 and A2 were not individual stimuli; rather, each comprised four pictures of faces, with each face belonging to a different person. The common feature of the faces in each category was the emotional expression, which was an angry expression in A1, and a happy one in A2. Perceptual classes were used to ensure that abstract stimuli would be equivalent to a particular emotional expression and not to idiosyncratic aspects of a particular face. Thus, the two 5-member equivalence classes included both perceptually related meaningful stimuli and arbitrarily related stimuli.

Visual inspection of the waveforms presented by both Barnes-Holmes et al. (2005) and Haimson et al. (2009) reveals that the greater negativity for the unrelated condition involving arbitrary stimuli started at around 400 ms. Haimson et al. (2009) also showed ERP waveforms for common English words. In this case, ERP waveforms indicate shorter latency for the N400 differentiation: starting at around 320 ms. Would the stimuli and parameters adopted in the current study produce equivalence relations even more similar to genuine semantic relations, at least when latency for the N400 differentiation is the criterion considered?

The primary purposes of the current study are, thus, to (1) verify if we obtain conventional N400 differentiation after the relational training (i.e., larger N400 responses for non-equivalent pairs than for equivalent ones), (2) contrast our data with those reported by Barnes-Holmes et al. (2005) and Haimson et al. (2009), and (3) provide a context for discussing the role that the ERP technique may play in stimulus equivalence research.

Method

Participants

Participants were 20 undergraduates, 6 males and 14 females, ranging in age from 18 to 26 years. All participants were students from a Brazilian university. Their native language was Portuguese, and they were not familiar with stimulus equivalence or related phenomena and concepts.

Equipment, setting, and stimuli

Sessions were conducted in a 3-m × 3-m laboratory room and were approximately 90 to 120 min long. An Apple Macintosh G4 microcomputer using the MTS software (Version 10.32; Dube & Hiris, 1997) was employed to generate equivalence classes. This equipment presented stimuli and recorded responses. Each trial displayed five white windows (6 cm × 6 cm) on a gray screen, one at the center and one near each of the monitor's corners. Participants responded by moving the computer's mouse to position a cursor on a window and then clicking the mouse's button. Another microcomputer using the Neuron-Spectron. NET software was employed for the "lexical decision" tasks. The equipment controlled both an EEG amplifier Neurosoft Neuron-Spectron 4/EPM and a computer monitor in whose screen visual stimuli were presented. Each lexical decision trial displayed a sequence of stimuli on a white screen. Participants responded by pressing two keys on the computer's keyboard: "S" and "N" (details below).

Figure 1 presents the stimuli employed in the experiment. Set A was comprised of eight pictures of human faces: four happy faces (A1) and four angry faces (A2). Sets B, C, D, and E were comprised of two arbitrary pictures each.

The pictures were extracted from the *Pictures of Facial Affect*© CD-ROM, purchased from Paul Ekman's website (www.paulekman.com). Several pictures of human faces depicting expressions of happiness, anger,

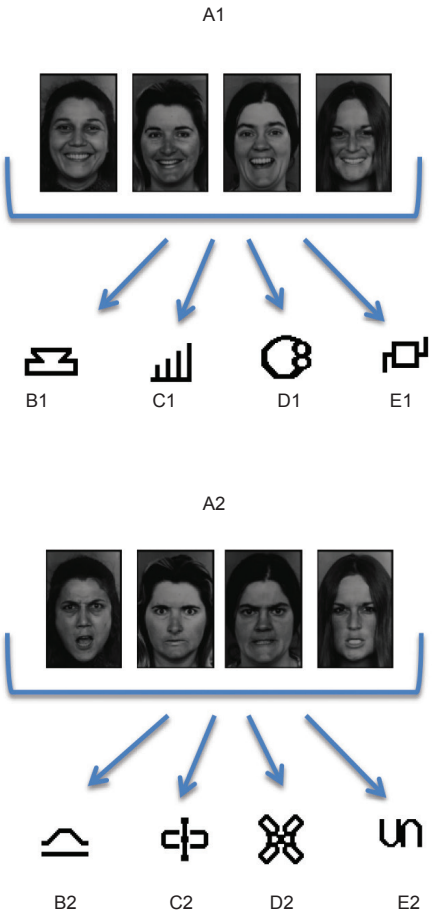

Figure 1. Schematic representation of the trained relations.

disgust, fear, surprise, and sadness are recorded on this CD-ROM. The pictures selected for this study were judged to be expressions of happiness and anger by 100% of the judges who viewed and rated the stimuli (Ekman & Friesen, 1976).

Procedure

Phase 1: Establishment of equivalence classes.

Each matching-to-sample trial began with the presentation of the sample stimulus in the central window. A click on the central window removed the sample and, after a delay of 2 s, produced the presentation of the two comparison stimuli, one in each of two of the peripheral windows. The two other peripheral windows remained blank. A click on the window containing the stimulus designated as correct produced a sequence of tones and a display of stars moving on the computer screen. Incorrect responses blackened the screen for 3 s. The consequence for a correct or an incorrect response ended the trial, and, after a 2-s intertrial interval, a new trial began.

Participants learned the conditional discrimination AB first, with a block of 24 AB trials in which samples A1 and A2 were presented 12 times each in a randomized sequence. Sample A1 could be any one of the happy faces and sample A2 could be any one of the angry faces. The positions of the comparison stimuli were determined according to a randomized sequence. In the first eight trials of this block a written prompt appeared on the screen. The Portuguese equivalent of the phrase "When this is here" appeared above the sample, and the Portuguese equivalent of "Choose this" appeared above the correct comparison. These eight trials were followed by 16 trials without these prompts. If the learning criterion (correct choices in all of the 24 trials) was not achieved, the block was repeated. AB teaching ended when this criterion was attained, and teaching of the AC relation began, using the same procedure. When the participant made correct choices in all AC trials, AD training started and, finally, AE training started with a similar procedure. Each of these blocks – AB, AC, AD, and AE – could be repeated for a maximum of three times. If the participant did not achieve criterion in three presentations of a block, she or he was dismissed.

The next block verified maintenance of the cumulative baseline–AB, AC, AD, and AE–and mixed 12 trials of each of these conditional relations, therefore comprising 48 trials in a randomized sequence. This block was repeated (for a maximum of three times) until the participant made no more than one incorrect selection.

When this criterion was achieved, the Portuguese equivalent of the message "The computer will no longer signal if your choices are correct or wrong" was displayed on the screen, and the cumulative baseline block was repeated without differential consequences for correct and incorrect responses until participants made no more than one error. If the participant made wrong choices in more than five trials, he or she returned to the cumulative baseline with differential consequences. Figure 1 shows a schematic representation of the trained relations in this experiment.

Table 1. Sequence of training and testing trial blocks presented to participants during Phase 1.

	BLOCK	NUMBER OF TRIALS
	with feedback	
1.	AB	24
2.	AC	24
3.	AD	24
4.	AE	24
5.	AB/AC/AD/AE	48
	without feedback	
6.	AB/AC/AD/AE	48
7.	Probe BC	16
8.	Probe CB	16
9.	AB/AC/AD/AE	48
10.	Probe BD	16
11.	Probe DB	16
12.	AB/AC/AD/AE	48
13.	Probe BE	16
14.	Probe EB	16

Equivalence probes. Six blocks of 16 probe trials without differential consequences tested 50% of the possible emergent relations between the arbitrary pictures that would document equivalence class formation. The first block evaluated the emergence of the BC derived relation; the second one tested the emergent conditional discrimination CB; the third block tested the derived relation BD; the forth block tested the relation DB; the fifth block tested the relation BE; finally, the sixth block tested the emergence of the EB relation. Between the second and the third probe blocks and also between the forth and the fifth probe blocks the cumulative baseline block without differential consequences was presented. Table 1 presents the sequence of trial blocks in this phase of the experiment.

The next phase was conducted only with participants who made no more than one error in each probe block. These participants met the criterion used to conclude that they formed happy and angry equivalence classes (i.e., one equivalence class containing the happy expression and four abstract stimuli and another equivalence class containing the angry expression and four abstract stimuli). The other students ended their participation.

Phase 2. Electrophysiological Testing

Electrophysiological recording. A microcomputer controlled by the Neuron-Spectron.NET software was used for presenting visual stimuli and recording responses on a keyboard. All electrophysiological activity was collected on a Neurosoft Neuron-Spectron 4/EPM Acquisition System (low frequency cutoff at 0.1 Hz; high frequency cutoff at 100 Hz).

EEG recordings were obtained from seven standardized locations (FZ, CZ, C3, C4, PZ, P3, and P4, according to the 10–20 system). Leads were individually attached to the participants' scalp and their impedances were less than 5 kΩ. All leads were referenced to linked mastoids. The continuous analog activity from all channels was synchronized with the presentation of the visual stimuli. The electrophysiological data were digitized at a rate of 1000 Hz and stored for subsequent processing.

Stimulus presentations. Prior to testing, participants were given 96 practice trials (48 with related and 48 with unrelated stimuli pairs whose equivalence status was evaluated in Phase 1) to familiarize them with the task. On each of 96 trials, the pair of stimuli was presented successively. The presenting monitor was located about 85 cm from the participant's head. The horizontal and vertical visual angles of each stimulus ranged from 4º and 6º. The first stimulus of each pair (termed ''the prime'') was presented for 600 ms. After 500 ms, the second stimulus (''the target'') was presented, also for 600 ms.

During practice, participants were instructed verbally to observe the stimulus presentations and to identify the pairs as related or unrelated. Their choice was made by pressing either of two buttons to indicate the relational status of each pair. No differential consequences followed responses. Intertrial intervals were 2 s, defined from the termination of the second stimulus of the pair. Participants were allowed a rest period of approximately 2 min after the first 48 pairs.

Testing trials involved only stimulus pairs whose equivalence status was not evaluated in Phase 1. Potentially related and unrelated stimulus pairs were presented on the computer screen. Faces were not used in this portion of the study. Potentially related pairs had been positive comparison stimuli selected in relation to the same emotional valence during the training. The unrelated pairs had been selected in relation to different valences. For all participants, 96 stimulus pairs were presented (48 related and 48 unrelated).

Participants viewed all stimuli from a distance of 85 cm. The first stimulus of each pair was presented at the center of a computer screen for 600 ms. After a 500-ms delay, the second stimulus appeared for 600 ms. After the second stimulus was removed, the screen was blank for 500 ms period and then a question mark was presented for 2000 ms as a response prompt. If the stimuli were members of the same class, the participant responded by pressing the "S" key on the computer keyboard; if they were not related, s/he pressed the "N" key. After a 2000 ms intertrial interval, a new trial began.

Before testing began, the experimenter instructed the participant to judge whether or not the two stimuli of a given pair were related. In addition, the participant was asked to refrain from blinking when the second stimulus of the pair was presented.

Subsequent to the electrophysiological procedure, all participants were debriefed and encouraged to ask questions or express concerns (if any).

Results

Participants took approximately 45-70 minutes to complete the equivalence training. Nineteen participants met the criterion used to conclude that they formed equivalence classes (no more than one error in each probe block). One participant did not reach this criterion and his electrophysiological data were not included in the following analysis.

All trials on which EEG activity exceeded ±75 mV were rejected (Silva-Pereyra et al., 1999). Data from another five participants whose electrophysiological data were affected by excessive artifacts–in more than 40% of the trials–were not included as well.

Event-related potentials were time-locked to the presentation of the target image in each trial, with one average generated for trials presenting stimuli from the same equivalence class and another generated for trials presenting stimuli from different equivalence classes. A baseline of 200 ms prior to stimulus onset was calculated and the epochs extended to 800 ms post-stimulus. The averaged EEG data were calculated across 14 participants who responded consistently with the formation of equivalence classes. The grand average waveforms for each of the seven electrode sites for equivalent and non-equivalent trial types are shown in Figure 2.

Visual inspection of these waveforms indicates that the most pronounced negative deflections occurred for the non-equivalent trial types, commencing at around 350 ms. It is important to note that Figure 2 is fairly representative of individual data: 12 from the 14 participants whose data were included in the grand average showed N400 responses for non-equivalent stimuli.

To detect reliable differences between the ERPs to equivalent and non-equivalent stimulus pairs, the ERPs from these conditions were submitted to a repeated measures, two-tailed permutation test based on the tmax statistic (Blair & Karniski, 1993). This permutation test analysis was used in lieu of more conventional mean amplitude ANOVAs because it provides better spatial and temporal resolution than conventional ANOVAs while maintaining a desired family-wise alpha level (i.e., it corrects for the large number of comparisons). Moreover, the tmax statistic was chosen for this permutation test because it has been shown to

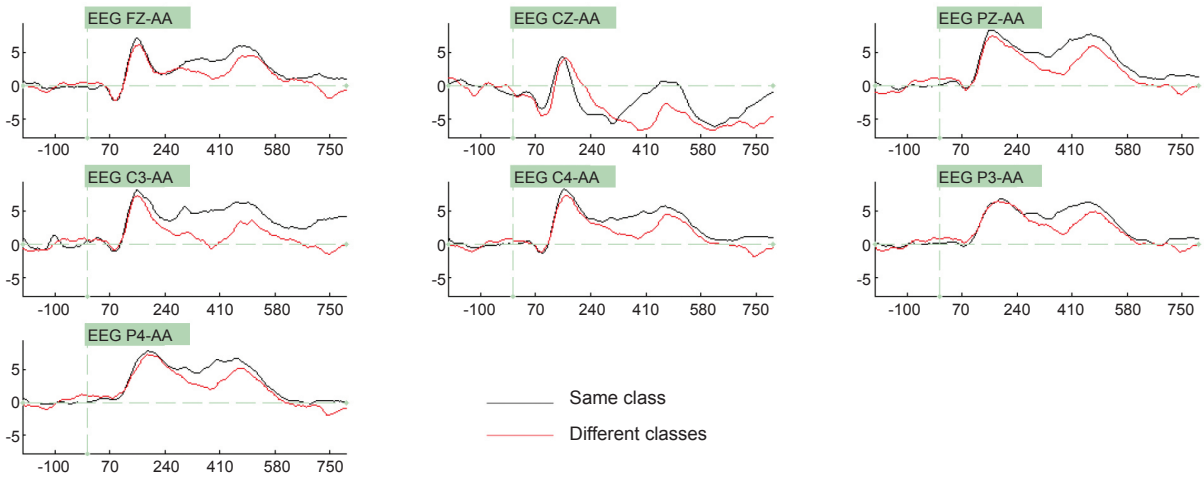

Figure 2. Waveforms obtained for pairs of equivalent stimuli (black lines) and non-equivalent stimuli (red lines) presented during the same/different tasks from participants who responded consistently with the formation of equivalence classes. Fz, Cz, Pz, C3, C4, P3, and P4 refer to locations where the electrodes were placed according to the international 10-20 system. A "z" (zero) refers to an electrode placed in the midline. F, C, and P refer to "frontal", "central" and "parietal", respectively. Even numbers refer to electrodes placed over the right hemisphere; odd numbers refer to electrodes placed on the left hemisphere. "AA" indicates that the electrodes were referenced to the average of the left and right mastoids. Negative voltage is plotted downwards.

have relatively good power for data (like ERPs) whose dimensions are highly correlated (Hemmelman et al., 2004). 2500 permutations were used to estimate the distribution of the null hypothesis as it is over twice the number recommend by Manly (1997) for a family-wise alpha level of 0.05. All time points between 100 and 900 ms at all 7 scalp electrodes were included in the test (i.e., 4900 total comparisons). Repeated measures t-tests were performed for each comparison using the original data and 2500 random within-participant permutations of the data using a family-wise alpha level of 0.05. The most extreme t-score in each of the 2501 sets of tests (i.e., the "tmax" of each set of tests) was recorded and used to estimate the tmax distribution of the null hypothesis (i.e., no difference between conditions[2]). Based on this estimate, critical t-scores of +/− 5.32 (df=13) were derived. In other words, any differences in the original data that exceeded a t-score of +/− 5.32 were deemed reliable.

Figure 3 presents a "raster" diagram of the permutation test.

Reliable differences between the ERPs to different-classes and same-classes stimuli were observed from 355 ms to 398 ms after the presentation of the target stimuli in the tasks in electrodes Cz, Pz, C3, P3, and P4.

Discussion

Electrodes placed over the central and parietal scalp regions registered larger N400 responses for non-equivalent stimulus pairs than for equivalent stimulus pairs. This result may be considered as a further indication that stimulus equivalence is an appropriate experimental model of semantic relations. In this section, we will present some differences between the current study and the two previous works that reported similar findings (Barnes-Holmes et al., 2005; Haimson et al., 2009). Moreover, we will discuss possible impacts of the experimental parameters adopted in this study on N400 latency and suggest that the ERP technique may provide a measure of semantic relations for research involving stimulus equivalence.

Differently from Barnes-Holmes et al. (2005), a single stimulus paradigm in which only one stimulus was presented at a time during the electrophysiological recording was employed. This single stimulus paradigm improved our ability to time-lock the ERPs to the instant that the target stimulus was shown. In addition, a single stimulus paradigm allowed us to establish a stable baseline prior to the presentation of the target stimulus that was identical across conditions. Non-stable baseline or different baselines across experimental conditions can be a major problem since the measured voltage in ERP experiments reflects the difference between the amplitude in the measured window and the amplitude in the prestimulus period (Luck, 2005). If prestimulus activity differs across conditions, any difference in measured amplitudes between conditions might reflect prestimulus differences rather than post-stimulus differences. In the current study, the baseline was stable and identical across conditions such that only the target stimuli would elicit a condition difference in the ERPs.

The present study did not include control presentations of related and unrelated real words during the electrophysiological recording. Thus, we cannot compare ERPs evoked by the intra-experimentally established relations and those acquired extra-

[2] More specifically, the null hypothesis of the permutation test is that positive differences between conditions could have just as likely been negative differences and vice-versa. Thus, the distribution of the null hypothesis is symmetric around a difference of 0.

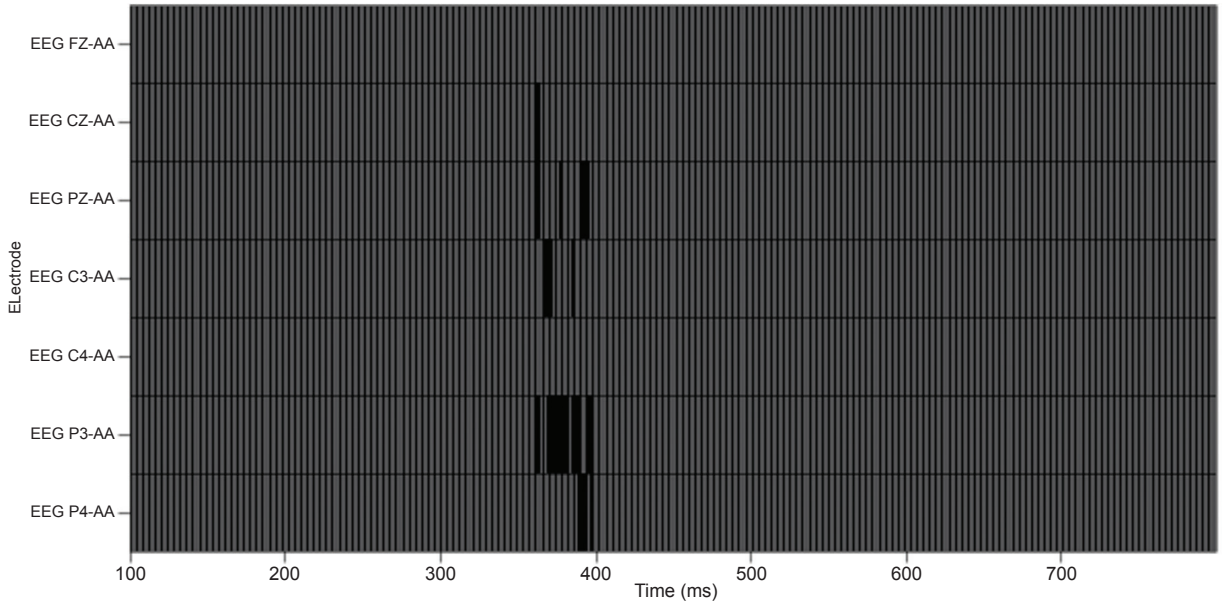

Figure 3. Raster diagram of the permutation test, with each electrode represented in a row and each time point represented by a column. Each box in the diagram therefore represents the result of a t-test. Black boxes indicate that the difference wave is significantly negative at that time point and electrode (even after effectively correcting for multiple comparisons). Gray boxes indicate that the t-test is not significant (after effectively correcting for multiple comparisons).

experimentally (via natural language processing) by our own participants. Nevertheless, it is still possible to contrast our ERP data with those presented in previous works that reported similar findings in order to discuss a possible role that the ERP technique may play in stimulus equivalence research.

An interesting difference emerges by contrasting our data with those reported by Barnes-Holmes et al. (2005) and Haimson et al. (2009) regarding the latency for the achievement of N400 responses. Visual inspection of the waveforms presented by both Barnes-Holmes et al. (2005, Fig. 1) and Haimson et al. (2009, Figs. 4 and 6) for intraexperimentally defined stimulus-stimulus relations reveals that the greater negativity for the unrelated condition started at around 400 ms. Haimson et al. (2009) also showed ERP waveforms for common words that were semantically related/unrelated in the native language of the participants. These waveforms indicate greater negativity for the unrelated condition starting around 320 ms. In the current study, most pronounced negative deflections occurred for the non-equivalent trial types at around 350 ms after the presentation of the target stimuli.

The N400 is a continuous measure whose latency is significantly modulated by three factors: language experience (e.g., Ardal, Donald, Meuter, Muldrew, Luce, 1990; Moreno & Kutas, 2005), age (e.g., Holcomb, Coffey, & Neville, 1992, Kutas & Iragui, 1998), and neurological disorders (e.g., Grillon, Ameli, & Glazer, 1991; Iragui, Kutas, & Salmon, 1996). The N400 is observed later in children (Holcomb, Coffey, & Neville, 1992), and its latency decreases with age and language experience to reach a minimum in early adulthood. In adulthood, language proficiency continues to impact N400 latency. As an example, Moreno and Kutas (2005) showed that, in bilingual individuals, N400 responses differentiate later for the nondominant as compared to the dominant language. Considering that all participants from the current study and from the former ones (Barnes-Holmes et al. 2005; Haimson et al., 2009) were young adults, differences between the N400 latencies might be related to the stimuli and experimental parameters employed to generate equivalence classes.

Differently from the relational training employed in the previous works, we used a sample-as-node training structure with a 2-s delay between the removal of the sample and the presentation of the comparison stimuli in each trial. As presented earlier, these parameters can improve the establishment of equivalence relations and the strength of the relations established. In addition, each equivalence class included both perceptually related meaningful stimuli and arbitrarily related stimuli. Fields, Arntzen, Nartey, and Eilefsen (2012) showed that equivalence class formation is more likely when training includes meaningful and arbitrary stimuli, in comparison with training only with arbitrary stimuli. Furthermore, equivalence classes that include meaningful stimuli share more features of natural language classes than classes comprising only arbitrary stimuli. It is possible that these experimental parameters and stimuli may have improved the participants' "proficiency" in the classes established, impacting the latency of the N400 component, which is a measure sensitive to linguistic proficiency.

Bortoloti and de Rose (2011b) argued that relational strength among equivalent stimuli and linguistic proficiency can vary in a similar fashion. The learning of a foreign language can exemplify this process. A word in

a person's first language may be very strongly related to a particular meaning. When learning a second language, (approximately) the same meaning might be related to a word in the second language, but the relation would not be, at least in the initial stages of learning, as strong as that with the word in the first language. We believe that the same rationale is applicable to equivalence classes: individuals could become more proficient in some equivalence relations than in others. If this is true, furthering the understanding of the differences in relational strength among equivalent stimuli will help to improve the equivalence paradigm, approaching the experimental model to relations involving symbols and meanings in the natural language.

The precise parameters that impact the relational strength among equivalent stimuli have not been fully explored, nor has this circumstance been generally recognized by researchers in this area. At present, just a few experimental conditions have been shown as potentially responsible for different levels of relational strength among equivalent stimuli. Delayed matching to sample and the involvement of emotional meaningful stimuli have been shown as responsible for strengthening relations among equivalent stimuli (Bortoloti & de Rose, 2009, 2011b, 2012); overtraining of baseline relations also strengthened emergent equivalence relations (Bortoloti, Rodrigues, Cortez, Pimentel, & de Rose, 2013). On the other hand, nodal distance (e.g., Bortoloti & de Rose, 2009; Fields et al., 2005) and number of logical properties involved in the relation (Doran & Fields, 2012) have been shown as responsible for weakening relations among equivalent stimuli. A major reason for this low number of studies, we believe, lies on difficulties to demonstrate different levels of relational strength through the standard procedures of the stimulus equivalence paradigm, based on matching to sample trials.

Even considering relations extraexperimentally established by bilingual individuals, matching to sample trials involving words and pictures would not detect differences between words from a person's first and second languages due to the all or none nature of this procedure. Based on the matching to sample trials, we would conclude that words in the second language are equivalent to the pictures. This procedure would capture an important property of the second language words–the pictures and the words are related by meaning–, but would not give a complete account. A quantitative dimension would have been left out. The same would occur with arbitrary equivalence relations defined intraexperimentally. Relations between stimuli in an equivalence class would be graded but these quantitative differences cannot be captured in forced-choice matching probes.

To access the strength of an equivalence relation, it is important to include post-class-formation tests based on a continuous measure in addition to the matching to sample procedure. ERPs such as the N400 component are continuous measures that may permit an assessment of the degree of relatedness between stimuli. The results of this study confirm that the N400 technique is useful for assessing the presence of equivalence relations and possibly the degree of strength of those relations, suggesting that behavioral analysis of symbolic processes may benefit from companion studies using electrophysiology. With the ERP technique, events within the brain become available and can be integrated to external events as a unit of analysis able to deal with features of symbolic relations that matching-to-sample procedures would leave out.

Acknowledgements

Renato Bortoloti was supported by the São Paulo Research Foundation (FAPESP) through a young investigator grant (#2011/50561-3). Naiene Pimentel was supported by a post-doctoral fellowship from Coordenção de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Julio C. de Rose has a research productivity grant from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). This study was part of the research program of Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino (Deisy G. de Souza, chairperson), supported by grants from CNPq (Grant #573972/2008-7) and FAPESP (Grant #08/57705-8). We thank Deisy de Souza for her continuing support of this project.

References

Ardal, S., Donald, M. W., Meuter, R., Muldrew, S., & Luce, M. (1990). Brain responses to semantic incongruity in bilinguals. *Brain and Language, 39*, 187-205.

Arntzen, E. (2006). Delayed matching to sample and stimulus equivalence: probability of responding in accord with equivalence as a function of different delays. *The Psychological Record*, *56*, 135-167.

Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training structure. *The Psychological Record, 47*, 309-320.

Arntzen, E, & Lian, T. (2010). Trained and derived relations with pictures versus abstract stimuli as nodes. *The Psychological Record*, *60*, 659–678.

Barnes-Holmes, D., Keane, J., Barnes-Holmes, Y., & Smeets, P. M. (2000). A derived transfer of emotive functions as a means of establishing differential preferences for soft drinks. *The Psychological Record, 50*, 493-511.

Barnes-Holmes, D., Staunton, C., Whelan, R., Barnes-Holmes, Y., Commins, S., Walsh, D., Stewart, I., Smeets, P. M., & Dymond, S. (2005). Derived stimulus relations, semantic priming, and event-related potentials: testing a behavioral theory of semantic networks. *Journal of the Experimental Analysis of Behavior*, *84*, 417-433.

Blair, R. C., & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology*, 30(5), 518-524.

Bornkessel, I., Schlesewsky, M., & Friederici, A.D. (2002). Grammar overrides frequency: evidence from the online processing of flexible word order. *Cognition, 85*, B21-B30.

Bortoloti, R., & de Rose, J. C. (2009). Assessment of the relatedness of equivalent stimuli through a semantic differential. *The Psychological Record, 59,* 563-590.

Bortoloti, R. & de Rose, J. C. (2011a). Avaliação do efeito de dica semântica e da indução de significado entre estímulos abstratos equivalentes. *Psicologia: Reflexão e Crítica, 24 (2),* 1-13.

Bortoloti, R., & de Rose, J. C. (2011b). An "Orwellian" account of stimulus equivalence. Are some stimuli "more equivalent" than others? *European Journal of Behavior Analysis, 12*, 121-134.

Bortoloti, R., & de Rose, J. C. (2012). Equivalent stimuli are more strongly related after training with delayed than with simultaneous matching: a study using the Implicit Relational Assessment Procedure (IRAP). *The Psychological Record, 62,* 41-54.

Bortoloti, R., Rodrigues, N. C., Cortez, M. D., Pimentel, N., & de Rose, J. C. (2013). Overtraining increases the strength of equivalence relations. *Psychology & Neuroscience, 6*, 357-364. doi: 10.3922/j.psns.2013.3.13

Bush, K. M., Sidman, M., & de Rose, T. (1989). Contextual controlo f emergent equivalence relations. *Journal of the Experimental Analysis of Behavior, 51,* 29-45.

de Rose, J. C., McIlvane, W. J., Dube, W. V., Galpin, V. X., & Stoddard, L. T. (1988). Emergent simple discrimination established by indirect relation to differential consequences. *Journal of the Experimental Analysis of Behavior, 50,* 1-20.

Doran & Fields (2012). All stimuli are equal, but some are more equal than others: measuring relational preferences within an equivalence class. *Journal of the Experimental Analysis of Behavior, 98*, 243–256

Dougher, M. J., Augustson, E., Markham, M. R., Greenway, D. E., & Wulfert, E. (1994). The transfer of respondent eliciting and extinction functions through stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior, 50,* 1-20.

Dube, W., & Hiris, J. (1997). Matching to Sample Program (Version 11.08) [Computer software]. Waltham, MA: E. K. Shriver Center for Mental Retardation.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Fields, L., Reeve, K.F., Matneja, P., Varelas, A., Belanich, J., Fitzer, A, & Shamoun, K. (2002). The formation of generalized categorization repertoires: Effect of multiple domain, sample, and comparison training. *Journal of the Experimental Analysis of Behavior, 78*, 291-313.

Fields, Arntzen, Nartey, and Eilefsen (2012). Effects of a meaningful, a discriminative, and a meaningless stimulus on equivalence class formation. *Journal of the Experimental Analysis of Behavior*, 97, 163–181.

Fields, L., & Moss, P. (2008). The formation of partially elaborated and fully elaborated generalized equivalence classes. *Journal of the Experimental Analysis of Behavior, 90*, 135-168.

Grillon C., Ameli R., & Glazer W.M. (1991). N400 and semantic categorization in schizophrenia. *Biological Psychiatry*, *29*, 467-80.

Haimson, B., Wilkinson, K.M., Rosenquist, C., Ouimet, C. & McIlvane, W.J. (2009). Electrophysiological correlates of stimulus equivalence processes. *Journal of the Experimental Analysis of Behavior, 92*, 245-256.

Hayes, S. C., Kohlenberg, B. S., & Hayes, L. J. (1991). The transfer of contextual control over equivalence classes through equivalence classes: a possible model of social stereotyping. *Journal of the Experimental Analysis of Behavior, 56*, 505-518.

Hemmelmann, C., Horn, M., Reiterer, S., Schack, B., Susse, T., & Weiss, S. (2004). Multivariate tests for the evaluation of high-dimensional EEG data. *Journal of Neuroscience Methods*, 139(1), 111-120.

Holcomb, P.J., Coffey, S. A, & Neville, H. J. (1992). Visual and auditory sentence processing: a developmental analysis using event-related brain potentials. *Developmental Neuropsychology, 8*, 203-241.

Iragui, V., Kutas, M., & Salmon, D. P. (1996). Event-related potentials during semantic categorization in normal aging and senile dementia of the Alzheimer's type. *Electroencephalography and Clinical Neurophysiology*, *100*, 392-406.

Kutas, M. & Federmeier, K.D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*, 463-470.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, *207* (1), 203-205.

Kutas, M., & Iragui, V. (1998). The N400 in a semantic categorization task across six decades. *Electroencephalography and Clinical Neurophysiology*, 108, 456-471.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: The MIT Press.

Manly, B. F. J. (1997). *Randomization, Bootstrap, and Monte Carlo Methods in Biology* (2nd ed.). London: Chapman & Hall.

Moreno, E., & Kutas, M. (2005). Processing semantic anomalies in two languages: an electrophysiological exploration in both languages of Spanish-English bilinguals. *Brain Research: Cognitive Brain Research*, 22(2), 205-220.

Niedeggen, M. & Rösler, F. (1996). N400 effects related to incongruities in mental calculation problems. *Psychophysiology, 33*, S65.

Osterhout, L. (1997). On the brain responses to syntactic anomalies: manipulations of word position and word class reveal individual differences. *Brain and Language, 59*, 494-522.

O'Toole, C., Barnes-Holmes, D., & Smyth, S. (2007). A derived transfer of functions and the implicit association test. *Journal of the Experimental Analysis of Behavior*, *88*, 263–283.

Saunders, R. R., Saunders, K. J., Kirby, K. C., & Spradlin, J. E. (1988). The merger and development of equivalence classes by unreinforced conditional selection of comparison stimuli. *Journal of the Experimental Analysis of Behavior*, *50*, 145-162.

Sidman, M. (1994). *Equivalence relations and behavior. A research story.* Boston: Authors Cooperative.

Sidman, M., & Tailby, W. (1982). Conditional discrimination VS. Matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37,* 5-22.

Silva-Pereyra, J., Harmony, T., Villanuera, G., Fernandez,T., Rodriguez, M., Galan, L., et al. (1999). The N400 and lexical decisions: automatic or controlled processing? Clinical Neurophysiology, 40, 813–824.

Steinhauer, K. & Connolly, J. F. (2008). Event-related potentials in the study of language. In: Stemmer, B. & Whitaker, H. (eds.). *Handbook of the Cognitive Neuroscience of Language*. New York: Elsevier.

Vaidya, M., & Smith, K. N. (2006). Brief report: Delayed matching-to-sample training facilitates derived relational responding. *Experimental Analysis of Human Behavior Bulletin*, *24*, 9-16.