Available online at www.sciencedirect.com
SCIENCE DIRECT®
Brain and Language 89 (2004) 253–260
Brain and Language
www.elsevier.com/locate/b&l

Short Communication

Neural correlates of bimodal speech and gesture comprehension☆

Spencer D. Kelly,* Corinne Kravitz, and Michael Hopkins

Neuroscience Program, Psychology Department, Colgate University, 13346-1338 Hamilton, NY, USA

Accepted 25 August 2003

Abstract

The present study examined the neural correlates of speech and hand gesture comprehension in a naturalistic context. Fifteen participants watched audiovisual segments of speech and gesture while event-related potentials (ERPs) were recorded to the speech. Gesture influenced the ERPs to the speech. Specifically, there was a right-lateralized N400 effect—reflecting semantic integration—when gestures mismatched versus matched the speech. In addition, early sensory components in bilateral occipital and frontal sites differentiated speech accompanied by matching versus non-matching gestures. These results suggest that hand gestures may be integrated with speech at early and late stages of language processing.

1. Introduction

People of all ages, cultures, and backgrounds gesture when they speak. What function do these hand movements serve? Although there is consensus that gesture plays an important role during language production (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Hadar, Wenkert-Olenik, Krauss, & Soroker, 1998; Iverson & Goldin-Meadow, 1998; Kita & Özyürek, 2003; Krauss, 1998; McNeill, 1992; Morrel-Samuels & Krauss, 1992; Rauscher, Krauss, & Chen, 1996), there is considerable debate on whether gesture also serves a *communicative* function during language comprehension (Clark, 1996; Kelly, Barr, Church, & Lynch, 1999; Krauss, 1998; Krauss, Morrel-Samuels, & Colasante, 1991; McNeill, 1992).

On one side, researchers who take a "gesture as communication" stance contend that gesture and speech are tightly integrated, with gesture influencing speech processing even at the earliest stages of comprehension (Cassell, McNeill, & McCullough, 1999; Kelly et al., 1999; McNeill, 1992). The standard paradigm used to support this claim has been to present people with verbal and gestural communication and to demonstrate that gesture is closely and unconsciously integrated with speech at comprehension.

On the other side, researchers who take a "gesture as non-communication" stance argue that gesture and speech are independent systems and that gesture does not influence language comprehension in a significant way (Krauss, 1998; Krauss et al., 1991). These researchers argue that the "gesture as communication" studies have demonstrated, at best, a trivial relationship between speech and gesture at comprehension. For example, gesture may be used as "add-on" information to comprehend a communicator's meaning only *after* the speech has been processed. In other words, in rare cases when gesture does influence comprehension, the attention to gesture is post hoc and does not impact early stages of speech processing.

At the core of this debate is the fact that all of the previous studies have relied on indirect behavioral measures that do not provide access to the underlying neurocognitive processing of speech and gesture. The present experiment addresses this issue by using a more direct measure of the neurocognitive processing of speech and gesture: event-related potentials (ERPs). Because ERPs provide excellent temporal information—they indicate *when* neurocognitive processes occur—the ERP technique is perfectly suited for investigating the relationship of gesture and speech in language comprehension.

☆ The authors thank Dennis Molfese, Susan Goldin-Meadow, and Howard Nusbaum for making comments on earlier versions of this paper. We extend special gratitude to Ron Crans for programming the stimulus presentation software.
* Corresponding author. Fax: 1-315-228-7942.
E-mail address: skelly@mail.colgate.edu (S.D. Kelly).

0093-934X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S0093-934X(03)00335-3

The experiment addresses two specific questions. The first question asks whether gesture influences ERPs to speech? A strong version of the "gesture as non-communication" view predicts that because gesture is completely independent from language, gesture should not have any impact on the neural processing of the accompanying speech. In contrast, the "gesture as communication" view predicts that because gesture and speech are fundamentally integrated, the ERP data will support past behavioral research demonstrating that gestures do influence speech processing.

The second question asks if gesture *does* impact speech processing, what is the time-course of this influence? A weak form of the "gesture as non-communication" view predicts that a gesture influence could occur only after the brain had already processed the semantic content of the speech. For example, this view would predict a late ERP effect resembling an N400 (Kutas & Hillyard, 1984). Previous researchers have argued that this type of effect reflects post-semantic processing of speech (Holcomb, 1993; Osterhout & Holcomb, 1995). In contrast, the "gesture as communication" view predicts that gesture would also affect pre-semantic processing of the speech. For example, this view predicts that gesture would influence early portions of the brainwave—such as, sensory, P1–N1, and P2 components—that reflect low-level, sensory/phonological processing of speech (Rugg & Coles, 1995).

2. Methods

2.1. Participants

Fifteen right-handed, Caucasian college undergraduates (6 males, 9 females; mean age: 20) participated for course credit.

2.2. Materials

Participants watched digitized videos (created with a Sony DV100 digital camcorder and edited with iMovie Macintosh software) of a person producing audiovisual speech and gesture. The audiovisual segments contained a male actor (face and hands toward the camera) sitting at a table behind a tall, thin glass and a short, wide dish. Note that the actor's mouth, face, and hands were in full view during the audiovisual presentation of the speech and gesture. These stimuli have been used successfully in behavioral studies of speech and gesture comprehension (Kelly & Church, 1998). In each clip, the actor uttered one of four speech tokens that corresponded to a salient dimension of the objects: *tall* and *thin* (the glass) and *short* and *wide* (the dish). Importantly, these words were digitized and inserted into the video, so that the words were identical across all of the following gesture conditions. Each condition corresponded to a different relationship that gesture can have with speech as described in the literature (for more information on the various relationships between speech and gesture, see McNeill, 1992).

One relationship is a gesture that conveys the same information as speech. These gestures made up the *matching* condition, in which the actor gestured to the same object and same dimension described in speech (e.g., said tall and gestured to the tallness of the tall, thin glass). A second relationship is a gesture that conveys similar *and* different information as the speech. These gestures comprised the *complementary* condition, in which the actor gestured to the same object described in speech, but represented a different dimension that complemented the speech (e.g., said tall but gestured to the *thinness* of the tall, thin glass). A third relationship is a gesture that conveys primarily different information from the speech. These gestures were in the *mismatching* condition, in which the actor gestured to the opposite object and also represented a dimension that contained different information as the speech (e.g., said tall and gestured to the *shortness* of the short, wide *dish*). Note that although the gestures in both the *complementary* and *mismatching* stimuli conveyed different *representational* content than the accompanying speech, the mismatching stimuli provided different *indexical* information as well. Finally, a fourth "relationship" is speech with no gesture. These gestures made up the *no gesture* condition and served as a baseline of speech comprehension without gesture.

The rationale for these conditions is threefold. First, the *no gesture* condition compared to the three gesture conditions will reveal differences in the brain's processing of speech with and without gesture. This comparison addresses the first question concerning whether gesture influences ERPs to speech. Second, differences between the *matching* versus *complementary* and *mismatching* conditions will reveal the brain's processing of speech when the *content* of the gesture conveys the same versus different information as the speech. This comparison addresses the issue of whether gesture influences speech due to mere "hand waving" or whether the substance of gesture actually matters. And third, differences between *complementary* and *mismatching* conditions will determine how much the semantic difference between speech and gesture influences language processing. That is, do complementary gestures that differ only in *representational* content from speech influence language processing as extensively as mismatching gestures that differ in *representational* and *indexical* content? These final two comparisons address the second question of when the content of gesture influences the brain's processing of speech.

There were 48 tokens of each condition, yielding a total of 192 trials. Each audiovisual segment was 3 s in duration, with the first portion showing a static image of the actor sitting behind the objects facing the camera. In all conditions, speech onset occurred at 2 s. In the gesture conditions, the onset of the gesture began 1200 ms after video onset and ended 1967 ms after onset (one frame before speech onset). The gesture was held in position during speech presentation and remained held until the video ended. The variable inter-stimulus interval was 1000–1500 ms.

2.3. Procedure

Participants were fitted with a 128-electrode Geodesic ERP net. The EEG was sampled at 250 Hz using a band pass filter of 0.1–30 Hz. The impedances were kept below 50 kΩ (the Geonet system uses high-impedance amplifiers). Eye artifacts were monitored with 4 EOG electrodes. All artifacts were corrected offline. Individual ERPs were segmented starting 100 ms before, and continuing 900 ms following stimulus onset.

During stimulus presentation, participants were instructed to press one key on a button box if the speech referred to the tall, thin glass and another key if the speech referred to the short, wide dish. The ERPs were taken to the speech stimuli.

2.4. Design and analysis

The experiment was a based on a 4 (Gesture condition) × 5 (Electrode site) × 2 (Hemisphere) within-subjects design with the auditory ERPs as the dependent measure. The Gesture manipulation refers to the four relationships of gesture to speech. The Electrode manipulation refers to the location of clusters of electrodes on the scalp. Based on previous studies, the 128 electrodes were broken up into five clusters of channels that corresponded roughly to basic anatomical structures of the brain. This electrode clustering technique has been successfully used in previous research to facilitate more manageable statistical analyses (Curran, 1999). The technique used in the present study is slightly different because it more closely follows natural anatomical divisions over bilateral frontal, parietal, central, temporal, and occipital brain regions. The Hemisphere manipulation refers to the fact that there were two sets of 5 electrode clusters, one over left hemisphere sites and one cluster over right hemisphere sites.

The data were analyzed using a temporal principal components analysis (PCA). This analysis is similar to a factor analysis, except that instead of uncovering correlations among data points, it identifies variability along the temporal dimension of the brainwave and assigns different factor scores and weights to those portions of the wave. These factors are blind to the conditions of the experiment and produce the same scores regardless of experimental manipulations. Although there has been debate about using this analysis technique (Wood & McCarthy, 1984), recent simulations have demonstrated that temporal PCA is as effective as traditional baseline-to-peak techniques in analyzing ERP data (Beauducel & Debener, 2003). The rationale for using this procedure is that it does not restrict the researcher's focus to only certain parts of the brainwave—rather, it identifies variability in many portions of the wave that may or may not be caused by the independent variables. This makes PCA well suited for analyzing data in an exploratory research area such as the present study.

3. Results and discussion

The PCA identified five factors that accounted for 87% of the total variance. Factor 1 accounted for 34% of the variance and ranged from 324 to 648 ms; factor 2 (25% of variance) ranged from 568 to 900 ms; factor 3 (13% of variance) ranged from 148 to 352 ms; factor 4 (8% of variance) ranged from 72 to 168 ms.; and factor 5 (7% of variance) ranged from 0 to 92 ms.

The weighted and Varimax-rotated PCA factors were submitted to a three-way repeated measures ANOVA with Gesture (no gesture, matching, complementary, and mismatching) × Electrode Cluster (frontal, central, parietal, temporal, and occipital) × Hemisphere (left and right) as the conditions. Dunn–Šidák *t* tests were used for all contrasts. Because we were only interested in main and interaction effects involving the Gesture condition, post hocs were run only on this condition. The ANOVAs on the five factors are organized into late and early effects in the brainwave, corresponding to the strong and weak versions of the "gesture as non-communication" views, respectively.

3.1. Late semantic components

The first two PCA factors reflected activity that occurred relatively late in the brain's processing of speech. These late components are associated with post-semantic processing and contextual integration (Rugg & Coles, 1995).

The ANOVA on PCA factor 2 (864 ms peak) revealed a significant main effect of Electrode ($F(4,56) = 10.39$, $p < .01$) but no effects for Gesture ($F(3,42) = 1.52$, ns) or Hemisphere ($F(1,14) = 0.09$, ns). In addition, there were no two-way interactions of Gesture by Electrode ($F(12,168) = 0.90$, ns), Gesture by Hemisphere ($F(3,42) = 0.67$, ns) or Electrode by Hemisphere ($F(4,56) = 0.19$, ns), nor was there a three-way interaction of Gesture by Electrode by Hemisphere ($F(12,168) = 0.84$, ns).

The ANOVA on PCA factor 1 (444 ms peak) yielded significant main effects of Gesture ($F(3,42) = 3.47$, $p < .05$), Electrode ($F(4,56) = 9.01$, $p < .01$), and

Hemisphere ($F(1,14) = 32.34$, $p < .01$). The main effect of Gesture was driven by a larger negativity across all electrode sites for the *no gesture* condition compared to *matching* ($tDS(3,42) = 2.71$, $p < .05$), *complementary* ($tDS(3,42) = 2.67$, $p < .05$), and *mismatching* conditions ($tDS(3,42) = 2.63$, $p < .05$). This provides initial evidence against the strong form of the "gesture as non-communication" view (i.e., gesture should not influence ERPs to speech) and suggests that the brain processed speech in different ways when speech was, versus was not, accompanied by gesture.

In addition, there was a significant Gesture by Electrode interaction ($F(12,168) = 10.05$, $p < .01$) and a marginal interaction of Electrode by Hemisphere ($F(4,56) = 2.90$, $p < .10$). However, there were no Gesture by Hemisphere ($F(3,42) = 1.75$, ns) or Gesture by Electrode by Hemisphere ($F(12,168) = 0.64$, ns) interactions.

The Gesture by Electrode interaction occurred at two electrode regions. In bilateral frontal regions, the *no gesture* stimuli produced a larger negativity than *matching* ($tDS(6,168) = 6.45$, $p < .01$), *complementary* ($tDS(6,168) = 6.00$, $p < .01$), and *mismatching* stimuli ($tDS(6,168) = 4.68$, $p < .01$). These results suggest that frontal sites were particularly sensitive to differentiate speech stimuli that were and were not accompanied by gesture.

Importantly, there were also effects within the gesture conditions. Specifically, at bilateral temporal sites, *mismatching* stimuli produced a larger negativity than *matching* stimuli ($tDS(6,168) = 2.98$, $p < .05$) but not *complementary* ($tDS(6,168) = 1.97$, ns) or *no gesture* ($tDS(6,168) = 1.90$, ns) stimuli (Fig. 1). On closer inspection, it became clear that this effect was driven largely by the right hemisphere sites: *mismatches* were significantly more negative than *matches* in the right hemisphere ($tDS(2,168) = 6.35$, $p < .01$) but not in the left hemisphere ($tDS(2,168) = 2.24$, ns).

This negativity bears resemblance to the N400 (Kutas & Hillyard, 1984) and reflects post-semantic processing of speech (Holcomb, 1993; Osterhout & Holcomb, 1995). Moreover, the fact that the effect was lateralized to the right hemisphere is consistent with past N400 findings (Kiefer, Weisbrod, Kern, Maier, & Spitzer, 1998; Kutas & Hillyard, 1982). Kiefer et al. (1998) explain that this lateralization occurs when words are preceded by indirectly related, but not directly related or non-related, semantic contexts. Indeed, the mismatching gestures conveyed information that was somewhat related to the speech (both the gesture and speech referred to dimensions of objects), but they were different enough (different dimensions and different objects) to be only indirectly related.

The results from factor 1 provide a clear answer to question one. Gesture does influence ERPs to speech. Although this rules out the strong version of the "gesture as non-communication" stance, it does not conflict with the weak version. Recall that the weak version predicts that gestures may have a post hoc impact on speech comprehension, affecting the comprehension process only *after* the brain has processed the semantic content of the speech. One might argue that the mismatching speech was processed initially without attention to the gesture, and the negativity at 400 ms (N400) reflected the brain's later attempt to integrate the lexical content of the speech with the previous gestural context (Holcomb, 1993; Osterhout & Holcomb, 1995). Therefore, the weak version cannot be ruled out without exploring earlier pre-semantic effects of gesture on speech comprehension.

3.2. Early pre-semantic components

The final three PCA factors reflected activity that occurred early in the brain's processing of speech. These

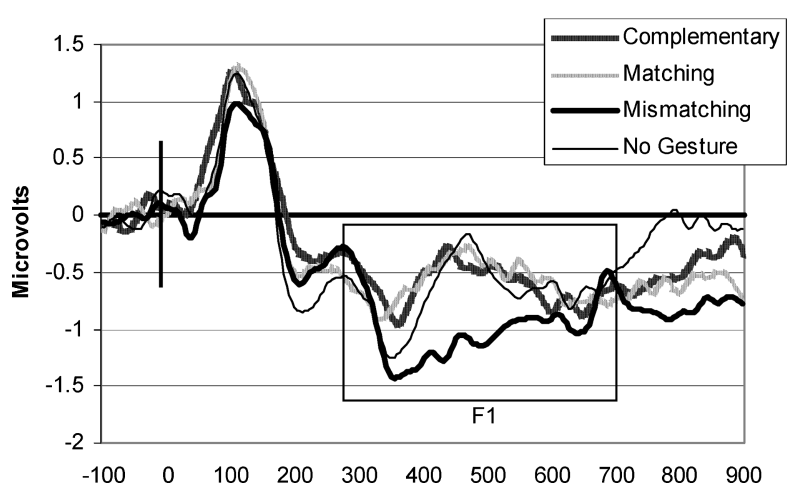

Fig. 1. The negativity from 324 to 648 ms (N400) at bilateral temporal sites for PCA factor 1. The *mismatching* condition produced a larger negativity than the *matching* condition, but it was not different than *complementary* and *no gesture* conditions.

early components are pre-semantic and reflect sensory and phonological processing (Rugg & Coles, 1995). If gestures influence speech at these stages of processing, it would be strong evidence against even the weak form of the "gesture as non-communication" view.

The ANOVA on PCA factor 3 (216 ms peak) revealed no significant effects for Gesture ($F(3,42) = 0.71$, ns) or Hemisphere ($F(1,14) = 0.06$, ns). In addition, there were no interactions of Gesture by Hemisphere ($F(3,42) = 1.11$, ns), Electrode by Hemisphere ($F(4,56) = 0.20$, ns) or Gesture by Electrode by Hemisphere ($F(12,168) = 0.57$, ns). However, there was a significant main effect of Electrode ($F(4,56) = 6.27$, $p < .01$) and a significant Gesture by Electrode interaction ($F(12,168) = 3.04$, $p < .05$). This interaction was driven by the *mismatch* condition producing a larger positivity at frontal sites than the *match* ($tDS(3,168) = 2.75$, $p < .05$) and *no gesture* ($tDS(3,168) = 4.11$, $p < .01$) conditions, but it was not different from the *complementary* condition ($tDS(3,168) = 1.70$, ns) (Fig. 2).

This effect is likely a P2 component and reflects phonological processing of the speech information (Dorman, 1974). This result is consistent with previous research demonstrating that incongruent non-verbal information (faces) produced a larger P2 to linguistic information (sentence fragments) than congruent non-verbal information (Pourtois, de Gelder, Vroomen, Rossion, & Crommelinck, 2000).

The ANOVA on PCA factor 4 (116 ms peak) revealed no significant effects for Gesture ($F(3,42) = 2.09$, ns) or Hemisphere ($F(1,14) = 0.01$, ns). In addition, there were no interactions of Gesture by Hemisphere ($F(3,42) = 0.34$, ns), Electrode by Hemisphere ($F(4,56) = 2.19$, ns) or Gesture by Electrode by Hemisphere ($F(12,168) = 0.97$, ns). However, there was a significant main effect of Electrode ($F(4,56) = 12.60$, $p < .01$) and a Gesture by Electrode interaction ($F(12,168) = 2.73$, $p < .05$). This interaction was driven by the *complementary* condition producing a larger positivity at frontal sites than the *match* ($tDS(2,168) = 2.51$, $p < .05$) and *no gesture* ($tDS(2,168) = 3.30$, $p < .01$) conditions, but there was no difference from the *mismatch* condition ($tDS(3,168) = 1.71$, ns) (Fig. 2). This effect corresponds to the P1–N1 component and likely reflects low-level auditory processing of the speech information (Näätänen & Picton, 1987). Previous research has demonstrated that this component reflects an early cross-modal influence of low-level visuospatial information on auditory processing (Eimer & Driver, 2001).

Finally, in order to determine whether gesture affected very early speech processing, we performed an ANOVA on PCA factor 5 (36 ms peak). This analysis revealed no significant effects for Gesture ($F(3,42) = 2.35$, ns), Electrode ($F(4,56) = 2.37$, ns) or Hemisphere ($F(1,14) = 1.89$, ns). In addition, there were no interactions of Gesture by Hemisphere ($F(3,42) = 0.04$, ns), Electrode by Hemisphere ($F(4,56) = 0.31$, ns) or Gesture by Electrode by Hemisphere ($F(12,168) = 1.40$, ns). However, there was a significant Gesture by Electrode interaction ($F(12,168) = 3.16$, $p < .05$). This interaction was driven by the *matching* condition producing a smaller negativity at occipital sites than the *complementary* ($tDS(3,168) = 3.00$, $p < .01$) and *mismatch* ($tDS(3,168) = 3.52$, $p < .01$) conditions, but it was not different from the *no gesture* condition ($tDS(3,168) = 0.76$, ns) (Fig. 3). This is strong evidence against the weak version of the "gesture as non-communication" argument, as gestures appear to influence how speech is acoustically encoded several hundred milliseconds prior to any semantic analysis of the speech content.

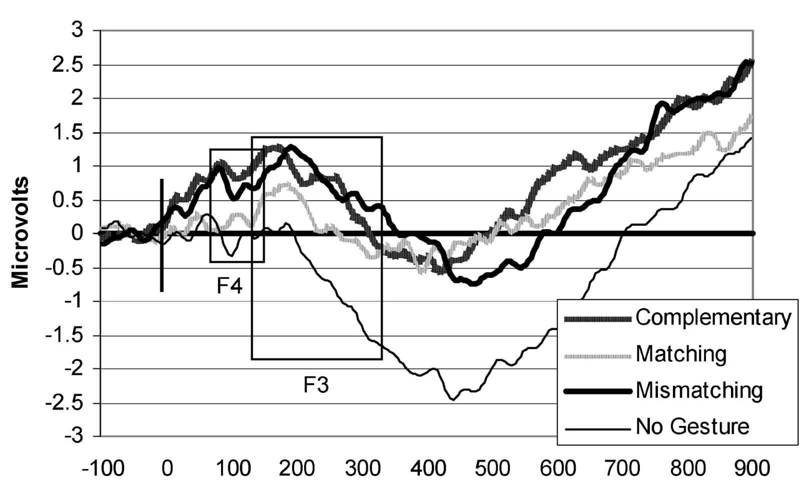

Fig. 2. P1–N1 and P2 effects at bilateral frontal sites for PCA factors 3 and 4. For PCA factor 4 (F4), the P1–N1 component was more positive for the *complementary* condition compared to the *matching* and *no gesture* conditions, but not the *mismatching* condition. For PCA factor 3 (F3), the P2 was more positive for *mismatching* condition compared to *matching* and *no gesture* conditions, but not the *complementary* condition.

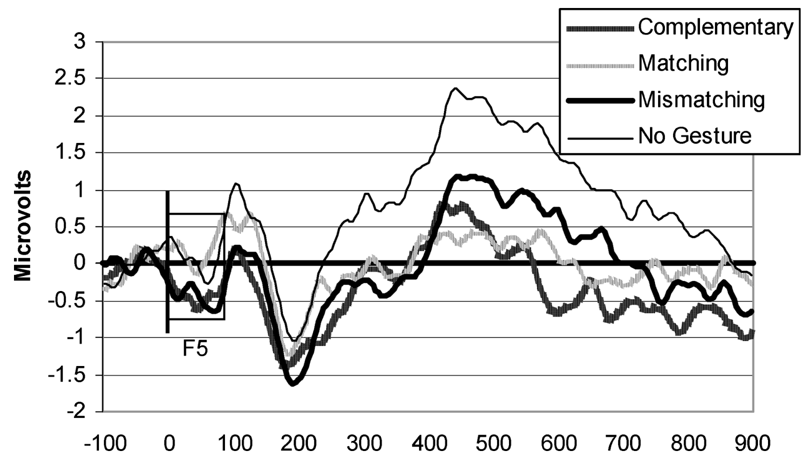

Fig. 3. The early sensory effect at bilateral occipital sites for PCA factor 5. The *complementary* and *mismatching* conditions were more negative than the *matching* condition, but not the *no gesture* condition.

4. General discussion

The results have provided answers to the two questions posed in this experiment, in addition to raising some interesting methodological considerations.

4.1. Does gesture impact ERPs to speech?

Hand gestures do impact the brain's processing of speech. The most basic effect was that there were different ERPs for speech that was accompanied by gesture compared to speech that was not. On the surface, this may seem trivial—the brain processes speech differently when visual motion precedes language than when it does not. However, closer analysis of the gesture conditions suggests that the *type* of visual motion makes a difference. Specifically, the brain produces different responses when gestures convey the same information as speech (*matching*) compared to when it produces different information (*complementary* and *mismatching*). So it is not mere hand waving—content matters.

It is interesting that *both* the *mismatching and complementary* stimuli were different than the *matching* stimuli for the early components. It may not be surprising that mismatching gestures produced a different speech response than matching gestures. After all, the mismatching gestures conveyed indexical information about a completely different object than the matching gestures. However, the *complementary* effect is more interesting. The only difference between the complementary and matching gestures was the *representational content* of the gestures. This finding suggests that communicators rely on the semantic content of speech *and* gesture when comprehending meaning. This result provides support for the "gesture as communication" view and suggests that gesture and speech may be tightly integrated at a deep semantic level (McNeill, 1992).

4.2. When does gesture influence ERPs to speech?

Gestures had different influences on speech across different portions of the brainwave. What might these different effects mean? The early effects (the early sensory component, P1–N1, and P2) revealed that the brain processed the speech in both the *mismatching* and *complementary* conditions differently than the *matching* condition. Although the present research is methodologically different from previous research demonstrating a low-level cross-modal influence of visuospatial information on early stages of auditory information (Eimer & Driver, 2001; Mangun & Hillyard, 1991), the present results suggest that the high-level visuospatial information conveyed through hand gestures may have an early cross-modal effect on speech processing. Specifically, gestures may create a visuospatial context that subsequently influences the sensory processing of the linguistic information that follows. This provides strong evidence that gesture and speech may be tightly integrated at even very early stages of processing (McNeill, 1992).

Interestingly, the later effects show a different pattern of results. Specifically, recall that *mismatching* stimuli produced a larger negativity at 400 ms than *matching* stimuli, but the *complementary* condition did not. Why would the *complementary* condition, which showed a sensory/phonological effect, not show a post-semantic effect? The answer may rest in the crucial difference between the *mismatching* and *complementary* conditions. Both conditions had gestures that conveyed different *representational* information than the speech, but only the mismatching stimuli presented different *indexical* information as well. This suggests that at late stages of processing, the semantic content of the complementary gestures were treated as partially consistent with the semantic content of the accompanying speech—indeed, the complementary gestures simply reflected a different dimension of the *same object* described in speech. Presented this way, one might not expect an N400 effect—which usually only occurs when there is semantic distance between items—for this condition. In contrast, the mismatching gestures conveyed indexical information about a *different object* than the speech. When the speech followed the mismatching gesture, it may have been treated as semantically more distant, and this greater distance may have caused the larger negativity at 400 ms (Kiefer et al., 1998).

In sum, gestures that convey subtly (*complementary*) and substantially (*mismatching*) different information from speech both appear to affect early stages of sensory/phonological processing of speech. However, only gestures that convey substantially different information from speech appear to affect later stages of post-semantic processing. This provides further support for the "gesture as communication" stance and suggests that gestures may be integrated with speech on many different levels of processing.

4.3. Methodological considerations

The present experiment makes an important methodological contribution to the established literature on the use of ERPs to study language processing. Previous research that has used naturalistic presentation of language stimuli (auditorily presented words in natural, connected speech) has discovered that context can influence speech processing very early (within the first 100 ms) in the brainwave (Holcomb & Neville, 1991). The present study went a step further and used an *audiovisual* presentation of language stimuli—a presentation that more closely approximates normal face-to-face communication than previous ERP studies on language processing. This mode of presentation helps explain the early sensory effect (PCA factor 5) found in the present study. For example, the effect makes sense when one considers that participants had access not only to auditory linguistic input, but also *visual* linguistic input. Because participants could see the actor on the stimulus tape produce the linguistic utterances, they had access to co-occurring information such as lip and mouth movements that naturally precede auditory onset of speech (Haxby, Hoffman, & Gobbini, 2002). These lip movements likely cued participants to linguistic information *before* they could actually *hear* it. This explains why gestures had such an early influence on speech in the brainwaves. Participants were probably using the complementary and mismatching gestures and the lip movements in tandem to anticipate speech onset. This makes the gesture effects potentially all the more interesting. Participants may have been using gesture to interpret lip movements of speech production milliseconds before the speech was ever heard!

5. Conclusion

Language naturally occurs in a rich communicative and audiovisual context. The results from the present study demonstrate that one aspect of this context—hand gesture—significantly impacts the comprehension of accompanying speech at multiple stages of language comprehension. This finding not only pertains to theories of gesture–speech integration, but also to any account of how the brain comprehends language in natural discourse.

References

Beauducel, A., & Debener, S. (2003). Misallocation of variance in event-related potentials: Simulation studies of the effects of test power, topography, and baseline-to-peak versus principal component quantifications. *Journal of Neuroscience Methods, 124*, 103–124.

Cassell, J., McNeill, D., & McCullough, K. E. (1999). Speech-gesture mismatches: Evidence for one underlying representation of linguistic and nonlinguistic information. *Pragmatics and Cognition, 7*(1), 1–34.

Clark, H. H. (1996). *Using language*. Cambridge, GB: Cambridge University Press.

Curran, T. (1999). The electrophysiology of incidental and intentional retrieval: ERP old/new effects in lexical decision and recognition memory. *Neuropsychologia, 37*(7), 771–785.

Dorman, M. F. (1974). Auditory evoked potential correlates of speech sound discrimination. *Perception and Psychophysics, 15*(2), 215–220.

Eimer, M., & Driver, J. (2001). Crossmodal links in endogenous and exogenous spatial attention: Evidence from event-related brain potential studies. *Neuroscience and Biobehavioral Reviews, 25*(6), 497–511.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesture lightens the load. *Psychological Science, 12*, 516–522.

Hadar, U., Wenkert-Olenik, D., Krauss, R. M., & Soroker, N. (1998). Gesture and the processing of speech: Neuropsychological evidence. *Brain & Language, 62*(1), 107–126.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry, 51*(1), 59–67.

Holcomb, P. (1993). Semantic priming and stimulus degradation: Implications for the role of N400 in language processing. *Psychophysiology, 30*(1), 47–61.

Holcomb, P. J., & Neville, H. J. (1991). Natural speech processing: An analysis using event-related brain potentials. *Psychobiology, 19*(4), 286–300.

Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature, 396*, 228.

Kelly, S. D., Barr, D., Church, R. B., & Lynch, K. (1999). Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of Memory and Language, 40*, 577–592.

Kelly, S. D., & Church, R. B. (1998). A comparison between children's and adults' ability to detect children's representational gestures. *Child Development, 69*, 85–93.

Kiefer, M., Weisbrod, M., Kern, I., Maier, S., & Spitzer, M. (1998). Right hemisphere activation during indirect semantic priming: Evidence from event-related potentials. *Brain and Language, 64*(3), 377–408.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language, 48*(1), 16–32.

Krauss, R. M. (1998). Why do we gesture when we speak? *Current Directions in Psychological Science, 7*(2), 54–60.

Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality & Social Psychology, 61*(5), 743–754.

Kutas, M., & Hillyard, S. A. (1982). The lateral distribution of event-related potentials during sentence processing. *Neuropsychologia, 20*(5), 579–590.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*(5947), 161–163.

Mangun, G. R., & Hillyard, S. A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *Journal of Experimental Psychology: Human Perception and Performance, 17*(4), 1057–1074.

McNeill, D. (1992). *Hand and mind: What gesture reveals about thought*. Chicago, IL: University of Chicago Press.

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18*(3), 615–622.

Näätänen, R., & Picton, T. W. (1987). The N1 wave of the human electric and magnetic response to sound: A review and analysis of the component structure. *Psychophysiology, 24*(4), 375–425.

Osterhout, L., & Holcomb, P. J. (1995). Event related potentials and language comprehension. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition* (pp. 171–215). London: Oxford University Press.

Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., & Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport: For Rapid Communication of Neuroscience Research, 11*(6), 1329–1333.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science, 7*(4), 226–231.

Rugg, M. D., & Coles, M. G. H. (1995). *Electrophysiology of mind: Event-related brain potentials and cognition*. London: Oxford University Press.

Wood, C. C., & McCarthy, G. (1984). Principal component analysis of event-related potentials: Simulation studies demonstrate misallocation of variance across components. *Electroencephalography Clinical Neurophysiology, 59*, 249–260.