Gesture – Speech Alignment in Deception

PhD project

Cristina Sánchez Conde
Universitat Pompeu Fabra

Supervisor: Dr. Pilar Prieto Vives

January 2017
Abstract

There is a considerable amount of research showing that deception affects nonverbal behavior (see Cohen et al. 2010), DePaulo, et al., 1985, Vrij, 2004, Vrij et al., 2000 for a review). However, very little work has been done on the temporal patterns of gesturer-speech integration in deception. The underlying hypothesis of this thesis will be that lying will induce greater degrees of gesture-speech asynchrony. It is well-known that gestures and speech are tightly aligned from a temporal point of view (see McNeill 1992). Specifically prominent parts of gestures (e.g., stroke, apex) tend to be temporally associated with prominent parts of speech (e.g., accented syllables, pitch peaks; see Esteve Gibert, & Prieto, P. 2013 for a review). The main objective of this PhD thesis will be to find out whether deception can significantly alter patterns of speech-gesture alignment and whether these are used in perception to detect deceiving speech. Two main studies will be conducted. The goal of the first study (in collaboration with Dr. Pereda’s Digital Humanities unit at Eurecat) will be to test whether deception affects the patterns of temporal alignment between deictic gestures and speech. Participants will be asked to collaborate in three deceit elicitation tasks. Results from two pilot experiments revealed that deceit items were produced with a longer lag between pitch peaks and the apex of the gesture, and apexes were shorter. However, we observed some individual differences that could be related to motivation and anxiety levels. Research shows that motivated deceivers tend to perform better with verbal cues, but worse with nonverbal cues (DePaulo et al. 1983) and we think that they might also show higher degrees of speech-gesture asynchrony. We would like to replicate the results of the first two tasks and correlate the degrees of gesture-speech asynchrony with motivation, arousal and anxiety measures. Thus a third experiment involving a deceit elicitation task will be carried out in which electrophysiological measures (Skin Conductance Response), which will be taken along with cognitive measures so as to observe the different degrees of asynchrony. The second study will focus on the effects of different magnitudes of gesture-speech asynchrony on the detection of deceit. To sum up, this project aims at assessing the role of gesture-speech integration both in the production and perception of lies and anxiety levels.

Keywords, deception, nonverbal behaviour, gesture-speech alignment, deictic gesture, pitch peak, Skin Conductance Response, stress and anxiety.
1. Introduction.

1.1. What is deception?

Lying is an act of everyday speech in normal social interactions. In this thesis, we will use the cover term deception, which covers a wide set of meanings. Buller et al. (2006), define it as "messages knowingly and intentionally transmitted to foster false beliefs or conclusions." According to Poggi et al. (2006) deceit can be described as “any act, or even non-act (omission) of an agent aimed at having another agent believe something false or not believe something true about the environment, the agent’s identity or its mental and emotional states.” Both definitions in fact encompass a variety of ways in which speakers can mislead their interlocutors: Lies, white lies, hyperbole, misdirection, equivocation, evasion, tall tales, half-truths, page hijacking, disinformation, omissions, concealment, strategic ambiguity, fakery, camouflage, mimicry, dazzling, exaggerations, decoys, phishing, spoofing, ruses, masking, feints and disguises (see Burgoon, 2013).

Within pragmatics, deception in conversation is typically used for benefit of the self and represents a covert violation of Grice’s cooperativeness principle (CP) and their four maxims. Deception typically violates one of the four Grice maxims, as follows (see Burgoon 2013 for more details). The maxim of quality is violated when speakers produce false utterances (fabrications) or utterances without evidence (with plausible stories that seem to be real). The maxim of quantity is violated when the contribution is not as informative as it is required (omitting details) or has much more detail to appear forthcoming. The maxim of relation is violated when the contribution is not relevant for the conversation or irrelevant information is intentionally included to go off the topic. Finally, the maxim of manner is violated when the contribution is intended to be obscure and ambiguous.

1.2. Why do people deceive?

Deception does not always occur for bad purposes. Burgoon (2013) claims that lies may be told to avoid conflict in situations where telling the truth may provoke conflict. One of the most common reasons why people tell lies for good purposes is politeness, maintaining other’s face or avoiding embarrassment. White lies or equivocation –not answering in a direct way- constitute usual types of deception used for this situation. But the word “deception” is more likely to be associated with negative conversational purposes, that is, when things are done to benefit the self either to gain a reward or to
escape from aversive consequences and affects negatively somebody else (Burgoon, 2013).

Some authors have argued that human behaviour is based on the following two main motivations: the Behavioral Approach System (BAS), which moves towards getting rewards, and the Behavioral Inhibition System (BIS), which moves towards the avoidance of punishment. The BIS/BAS Scales measure people’s sensitivity to both systems (Carver & White 1994). Several studies about deception have confirmed that liars who are motivated by a reward or those who avoid punishment perform better. (DePaulo et al. 1983; Ekman, 2009; Ekman & Frank, 1993; Vrij et al. 2000).

1.3. Anxiety and Cognitive Complexity

The motivation to lie or the fear to be caught while deceiving tends to increase anxiety levels. Lying is cognitively more complex than telling the truth. The Activation, Decision and Construction Model (ADCM) by Walczyk et al.(2012), Walczyk et al. (2009), Walczyk et al. (2003), and Walczyk et al. (2005) claims that three different phases can be found in deception, namely Activation, Decision and Construction. Activation is automatic: the truth travels from long-term memory to the working memory. Decision is intentional, that is, we decide whether telling the truth is convenient or not. If telling the truth involves negative consequences, we construct a lie in the third phase, which is also intentional (Masip & Herrero, 2015a). Decision and Construction both use cognitive resources that make increase the response latency –e.g., elapsed time between the end of a question and the beginning of the answer–.

Below I review a physiological measure which is related to stress and anxiety and which will be especially important in this thesis, namely Skin Conductance Response.

1.3.1. Skin Conductance Response (SCR)

The Skin Conductance Response (SCR), also called Electrodermal Response (EDR) or Galvanic Skin Response (GSR), is a measure of the electrical conductance of the skin and it varies depending on our stressed or relaxed state. The sympathetic nervous system controls sweat glands and responds to emotional and physical changes. Stress and anxiety raises skin conductance. The Skin Response is measured by electrodes on skin using Ohm's law principle that states skin resistance (R) is equal to the voltage (V) of two electrodes placed in skin and divided by the current (I) that passed through
the skin ($R=V/I$). Conductance is measured in microSiemens ($\mu$S) (Dawson et al. 2007).

To measure skin response, two electrodes or sensors are placed on the surface of the skin, typically from the phalanges. In an experiment, Scerbo et al. (1992) compared the SCR from distal and medial phalanges and results show that skin response amplitude is significantly higher measured in the distal phalanges. It was found that greater reactivity at the distal phalanges is correlated with greater number of sweat glands. For this reason, measurement is recommended to be done in distal site. (Freedman et al. 1994; Dawson et al. 2007).

Another important issue resides in the hand used for measurement. Dawson et al. (2007) report some studies claiming differences between left and right hand Skin response, but their interpretation seem to be unclear as it is difficult to make clear distinctions between emotional and non-emotional tasks for the experiments. SCR “in emotional tasks is presumably controlled primarily by the ipsilateral limbic system, whereas (…) in non-emotional tasks may be controlled by the contralateral system” (Dawson et al. 2007:163). Research on the topic examining SCR in both hands suggests including hand control although there is no definitive evidence about differences between both hands.

1.3.2. Skin Conductance Response (SCR) in Deception
The Guilty Knowledge Test (GKT) is a sort of test for detection of deception which uses Skin Conductance Responses. The GKT or CKT, Concealed Knowledge Test, registers skin response and other physiological responses at the same time subjects face multiple-choice questions. The correct answer is mixed between other alternatives for each question. The correct answer seems to be more psychologically significant to a deceiver or guilty subject than the other alternatives. For the truth teller (or innocent subject) all alternatives have the same significance. For that reason, deceivers presumably respond electrodermally speaking, more to the true answers (Lykken, 1959; Dawson et al. 2007). Using this test, Lykken (1981) found that nearly 90% of the time guilty subjects can be spotted and nearly 100% innocents are correctly classified. Even though Skin Conductance Responses have been used to detect deception, this measure is sensitive to a wide range of stimuli that makes unclear understand measures (Dawson et al. 2007). This means that while it is not possible to completely
identify stress or anxiety by Skin Conductance Response, it constitutes an interpretable measure in very controlled stimuli conditions where skin response occurs.

1.4. Detecting deception
There has been a longstanding scientific interest in the realization of deceptive speech and in developing techniques for detecting deception. Research to date has shown a good amount of speaker variability in the ways speakers express deception. Researchers have argued that there are no systematic behavioral indicators that consistently distinguish truth-telling vs. deceptive speech (Masip & Herrero, 2015ab). Some studies have singled out some personality factors favoring individual differences. For example, motivated liars or those who fear being caught perform differently from less motivated liars or those who do not feel fear. Ekman, names stakes the perceived consequences for successful and unsuccessful attempts at deception. The greater they are, the more they trigger detection apprehension. Detection apprehension may be higher when the stakes implicate avoiding a punishment or something unpleasant, as well as earning a reward (Ekman & Frank, 1993).

In the following two subsections I review some of the work that has been performed on the prosodic and gestural features of deceptive speech the last decades.

1.4.1. Gesture in Deception
One of the most prominent figures in the field of gestures, Paul Ekman, stated that not everybody behaves the same way when expressing themselves, and that gestures in deceptive speech reveal two different things, namely what people want to show, and what people want to hide. Lying requires making up an emotion in order to hide the one the liar is experiencing. Lies typically fail due to this reason and due to the failure of the liar preparing their line (Ekman & Frank, 1993). Lack of preparation may induce different clues to deceit: Fear, anxiety or annoyance signs in facial expressions as well as asymmetrical facial gestures. Spontaneous expressions are more or less symmetric in both sides of the face (Ekman, & Frank 1993; Ekman, 2009).

According to McNeill (1992), gesture is divided into five different phases: The Preparation Phase, which is the arm movement from a relaxed position to the gestural peak of effort. The Stroke or the peak of effort of the gesture, before and after the Apex. The Apex is the peak of maximum prominence of the gesture. And the last phase, Retraction or the movement of the arm from its highest extension to a relaxed
position. In Cohen et al. (2010), participants had to narrate a comic story twice: once telling the truth and once lying in three particular panels. Results showed fewer iconic gestures were produced in the deceptive condition. Moreover, gestures in deceptive speech had fewer post-stroke holds and shorter stroke phase durations than in those produced in true statements.

In Caso et al. (2006) participants were asked about a possession twice. In one of them they had to tell the truth, and in the other one, they had to lie. For the experiment, the suspicion factor was added, and the subjects were accused of lying in order to make them repeat their answers. Results showed a decreasing number of deictic gestures and self-adaptor gestures (touching face, biting lips, etc.), and an increase in metaphoric gestures (those used to represent abstract ideas) during deception before adding the suspicion factor. Results from the analysis of the answers with the suspicion factor showed an increase in metaphoric, rhythmic (those accompanying the prosody), and deictic gestures and a decrease in self-adaptor, emblematic (gestures with a specific meaning shared by a particular culture), and cohesive gestures (those connecting different parts of speech).

Burgoon et al. (2015) applied a software program called THEME to their investigation so as to analyze patterns of adaptor gestures, illustrators (those describing what the speaker says), and head movements. Results showed that deceivers perform fewer, shorter, and simpler gestural patterns than truthful responders. Deceivers tend to produce repetitive patterns and truth tellers are more eager to innovate and create new gestural patterns. These patterns are also less complex and are composed with fewer movements, which also make the gestures be shorter. The main reason for this to happen may be the increase of the cognitive load and the participants’ desire of their own control to appear credible.

### 1.4.2. Deceptive Speech

Studies about the characteristics of speech in deception seem to highlight some features, as well as contradictory results and individual differences in their implementation. Ekman (2009) pointed out that some speech characteristics can be considered helpful clues to spot deception. The author highlights numerous long and short pauses between words, mistakes and indirect speech and through roundabout expressions (e.g., periphrases).
Some empirical studies have focused on the occurrence of speech errors and hesitations in deceptive. Schober & Glick (2011) claim that a speech monitoring system is capable of detecting that in deceptive speech speakers show a higher number of speech errors that are words of the language errors than non-words speech errors. The speech monitoring system also detects that deceivers produce semantically related errors connected to the prior discourse and/or to the context. Taboo words are less likely to be produced as speech errors than non-taboo words. Vrij (2004) makes a distinction between speech errors (words, sentence change, repetitions, etc.) and speech hesitations (fillers such “ah”, “um”, etc.). The author points out that while some studies show both an increase of these errors and hesitations, in others the opposite occurs. The complexity of the lie seems to be the reason of such difference. When lies are complex, the number of errors increase, but when lies are easier, these number decrease.

A set of prosodic characteristics have been also pointed out as possible clues, such as fundamental frequency changes and speech rate, with contradictory results. In order to test whether the F0 changes occur in deception or not, Streeter et al. (1977) asked 32 men to answer five questions about politics, future, religion and values and to lie in two of them. The study concluded that F0 tends to rise in deceptive speech. By contrast, Villar et al. (2012) took into account the belief of the informants about speech changes in deception. The participants who thought F0 rises in deception showed a significantly higher in false statements. In contrast to this, the participants who though F0 does not rise and those who didn’t know whether if it rises or not, showed a little lower F0 in deceptive speech. However, Spence et al. (2012) found that in Italian there are no significant changes in F0 between true and false statements.

Regarding speech rate, some studies have found a greater number of words or a decreasing speech rate in false utterances. Ebesu & Miller (1994) compared four different situations, namely when the participants of an experiment tell the truth directly, when they just imply the truth, when they tell a lie directly and when they just imply the lie to another person. Results showed that deceivers perform with slower speech rates. Spence et al. (2012) found the same for Italian deceptive speech. Regarding the number of words, in Anolli & Ciceri (1997), different students were required to describe a picture three times: one telling the truth, one lying to an acquiescent judge, and another one lying to a suspicious judge. Results indicated the use of a greater number of words in the deceptive conditions.
By contrast, other studies have found non-significant results in the number of syllables and speech rate. In Hocking & Leathers (1980), participants were asked some questions about a video. In some of them, they were instructed to tell particular lies. Results showed that there is no significant difference between the speech rate of true and false statements. In another study, Feeley & DeTurck (1998) participants worked with three different conditions: truth, unsanctioned deception and sanctioned deception. They were given anagram tasks and then they were randomly assigned to one of the three conditions. Participants in the first condition had to tell the truth. Participants in the deceptive conditions had to cheat and look the answers of the anagram tasks. In the sanctioned condition, participants had to answer the interviewer lying about their high score in the tasks. In the unsanctioned condition, they were not given any instructions about the answer they had to perform. Results showed no significant differences between true and false answers in speech rate and in number of syllables. Spence et al. (2012) run an experiment where participants had to tell the truth and lie about two issues selected from a questionnaire for which they had strong feelings. They found that the number of words in the deceptive condition is lower than in the truth, yet not significant. The same happens to the acoustic duration.

1.5. Speech-Gesture synchrony

Even though research to date has shown that deception affects nonverbal behaviour (see Cohen et al. 2010; DePaulo et al. 1985; Vrij, 2004; Vrij et al. 2000) for a review). It is also well-known that gestures and speech are tightly aligned from a temporal point of view (see McNeill, 1992). Specifically prominent parts of gestures (e.g., stroke, apex) tend to be temporally associated with prominent parts of speech (e.g., accented syllables, pitch peaks; see Esteve-Gibert & Prieto, 2013 for a review).

In this section, we review the literature on gesture-speech temporal alignment.

Some investigations have shown that gestures are temporally synchronized with speech. McNeill (1992) stated that a message is coherent due to the integration of speech and gesture because both result from the same mental reaction. Temporal coherence takes place when both gesture and speech integrate into a single message, and interference in these temporal patterns might affect the production of the message. However, studies differ in which part of speech coordinates with which part of the
gesture. Butterworth & Beattie (1978) found that the prominent part of the gesture coordinates with the target or focused word, as they found gestures start before the correspondent target word. For Loehr (2004), the gesture apex is significantly aligned with pitch-accented syllables. In another experiment about deictic gestures in cartoon narrations, Nobe (1996) found that gesture coordinates with pitch peaks and intensity peaks. De Ruiter (1998) studied the alignment of the prosodic structures and pointing gestures in a contrastive production context. The author concluded the prosodic structure affect the gesture movement as the preparation and stroke phases before the apex were longer in a context of a phrase-final stress position. She found that the apex of deictic gestures aligned with accented syllables.

In general, all these investigations reveal that generally prominent syllables are aligned with the gesture apex or the gesture stroke. And specifically, some studies have found a close temporal relationship between the gesture apex or the gesture stroke and the F0 peak of the accented syllable. Chen et al. (2007) compared the temporal patterns of gesture strokes with fundamental frequency and intensity peak values. Results showed a high probability of alignment between gesture strokes and the pitch peak values. Similarly, Esteve-Gibert & Prieto (2013) run an experiment where participants had to read a target word with different stress patterns in contrastive focus while pointing at a screen. Results showed that the gestural apexes of the pointing gesture were temporally aligned with the peak of the fundamental frequency of the corresponding accented syllable.

---

1 Gesture phases (McNeill, 1992): Preparation phase (movement of the arm in a relaxed position until the gestural peak of effort), Stroke (peak of effort in the gesture) before the apex, Apex (peak of maximum prominence of the gesture), Stroke after the apex, and Retraction phase (arm movement from the highest extension to the relaxed position).

2 However, a few studies have found no such alignment between gestural and speech prominence. For example, Rusiewicz (2010) studied the relation between deictic gestures and contrastive pitch accents and syllable position. Participants had to tell the name of some objects in different pictures at the same time they pointed to them. Results showed that the gesture apexes were mostly aligned with the onset of the target word instead of being aligned with prominent syllables.
Morrel-Samuels et al. (1992) found that the beginning of the gesture occurs before the beginning of the speech and the interval is inversely proportional to the familiarity with the lexicon. This familiarity effect is also inversely proportional related to the duration of the gesture. Bergmann et al. (2011) concluded in their investigation that the more semantically related were speech and its associated gesture, the closer their temporal relation is.

However, to our knowledge no work has been carried out on the temporal patterns of gesture-speech integration in deception. To our knowledge, only Ekman (2009:144-145) has highlighted the importance of the duration of the gesture and how much time the gesture take to appear in relation to speech.

On the processing level, a study by Leonard & Cummins (2011) on another type of gestures, beat gestures\(^3\) found that listeners are sensitive to the synchronization between beat gestures and speech. Participants were exposed to gesture-speech combinations that were not synchronized temporally and they measured their reaction times on the stimuli. It was found, that participants were becoming aware of the lack of synchronization between gestures and speech on the 200 ms post stimulus onset. In sum, all these studies show that temporal synchronization of speech and gesture is an important characteristic of gesture-speech integration, and that prosody plays an important role in the temporal synchronization patterns.

2. **Goals of the thesis**

The underlying hypothesis of this thesis will be that lying will induce greater degrees of gesture-speech asynchrony. As we have seen, a variety of studies have confirmed the existence of a strong temporal relation between speech and gesture untruthful utterances. In particular, some of them corroborated that there is a sort of coordination between pitch peaks and the different phases of the deictic gesture. However, we don’t know what happens in deception. Does deception affect these gesture-speech temporal alignment patterns? As we see in the table below, the purpose of this work will to learn if this happens through the elaboration of two different studies.

\(^3\) Beat gestures do not carry any semantic content, but they go with the rhythm of the speech.
The main objective of the first study is to find out whether deception disrupts the temporal coordination between speech and deictic gestures. Our hypothesis is that deception will trigger asynchronic patterns between pitch peaks and the apexes of the deictic gesture. A second study will focus on perception. Its main goal is to know whether deception is perceived whenever there is asynchrony between pointing gestures and speech. Our hypothesis assumes people perceive there is deception in utterances when prominent positions in speech and the gesture strokes apexes of deictic gestures are not rightly aligned.

**Does deception affect gesture-speech temporal alignment?**

**Objectives:** The main objective of this PhD thesis will be to find out whether deception can (a) significantly alter patterns of speech-gesture alignment and (b) whether these are used in perception to detect deceiving speech.

**Hypotheses:** Deception affects gesture-speech temporal alignment

- **Study 1:**
  - Hypothesis: Deception causes asynchrony between both deictic gestural apexes and pitch peaks.
  - Objective: To find out whether there is temporal disruption between deictic gestures and speech in the deceptive condition.

- **Study 2:**
  - Hypothesis: People perceive deception when there is a temporal misalignment between deictic gestures and speech.
  - Objective: To test whether participants perceive deception when deictic gestures and speech are not temporally aligned.

3. **Theoretical Framework**

Many studies have been done about the realization of deceptive speech and the realization of gestures in deception, but research has revealed a large variability in how people deceive.

Paul Ekman (2009) has stated that gestures in deception reveal what people want to hide, but also what they want to show. The lack of preparation show different cues that
may be visible in many different ways. In deceptive speech, the number of gestures, particularly deictic gestures, decrease, but if the deceiver feels like a suspect, the number increases (Caso et al. 2006). We also know that gestures have fewer post-stroke holds and shorter stroke phases (see Cohen et al. 2010).

Research about speech in deception focuses in particular characteristics like the fundamental frequency or the number of words. Nevertheless, to our knowledge, there is still much contradiction between studies. Some studies claim that the fundamental frequency rise in deceptive speech (see Streeter et al. 1977) while others show that the fundamental frequency does not change in deception (see Spence et al. 2012; Villar et al. 2012). The same happens with the number of words. We can find studies claiming a greater number of words in deception (see Anolli & Ciceri, 1997) but others reveal a lower rate, yet not significant (see Spence et al. 2012).

However, a correlation between gesture and standard speech has been studied and can be used for deception studies. In 1992, McNeill explained that speech and gesture correlate in three different levels: 1) The phonological level which explains that the most prominent part of speech and the most prominent part of the gesture are aligned. 2) The semantic level which explains that gestures and speech occurring at the same time express the same idea. 3) The pragmatic level, which explains that speech and gesture occur for the same intentional function.

There are some studies that have proved the first correlation and confirm a gesture-speech temporal alignment. However, these investigations differ in how they align exactly. As we mentioned before, Butterworth & Beattie (1978), stated that the prominent part of the gesture coordinates with the target word. Also according to De Ruiter (1988) and Loehr (2004) specify that the gestural apex aligns with accented syllables. Other studies as in Chen et al. (2007) show that the stroke aligns with the pitch peak. More recently, Esteve-Gibert & Prieto (2013) revealed that the peak of the fundamental frequency of target accented syllables aligns with the apexes of the deictic gesture. This experiment was considered one of our bases for this project as they delimited their coding analysis with a clear methodology. They used the gesture phase coding from McNeill (1992) for the pointing gesture, and for speech, they delimited syllable boundaries and pitch peaks. Their results also revealed that gesture apexes and pitch peaks occur at the end of the syllable if the stress in non-final position and they are retracted when the stress is in final position.
However, as far as we know, there are not works done about the patterns of gesture-speech alignment in deception. This thesis main goal is to find out if deception can disrupt speech-gesture alignment and if this misalignment is perceived as deceptive speech. A better understanding about how gesture and speech correlates in deception will contribute not only to the linguistics field as it can be useful for practical purposes in the forensics field.

4. Experimental Studies

4.1. Study 1

Different studies reveal there is some kind of synchrony between accented syllables and strokes and apexes of pointing gestures. But, does deception affect the patterns of deictic gesture-speech alignment? For the first study, three different experiments will be conducted in order to test whether deception disrupts the synchrony between the apexes of deictic gestures and pitch peaks. The main differences between the three experiments are the complexity of the target sentences used in deception, as well as the topic about what people deceive and the degree of cognitive effort that the task needs. The study is being conducted in collaboration with Dr. Pereda’s Digital Humanities unit at Eurecat.

In the first experiment, participants will have to choose between two pictures lying where they will be asked to do so. The second experiment implies a higher cognitive load, as it requires an effort from the participants when they will have to tell actual lies about a particular story. In the third experiment, the subjects will have to lie about personal information. In the last experiment, we will check whether temporal misalignment occurs in the same way when participants deal with their own personal information.

Below we describe how the pilot experiments were run and discuss some pilot results with 5 participants. Even though they were run without skin conductance sensors, the final versions of the experiment will be conducted with those. Specifically, in the three experiments we will use Mindfield® eSence Skin Response, a small sensor which measures skin conductance through two electrodes attached to the index and middle fingers. The sensor will provide with feedback about galvanic skin response (GSV), also called electrodermal response (EDR) or skin conductance response (SCR), in μSiemens (μS). The electrodes of the sensor will be attached to the microphone or
headphone input of a smartphone in order to run the Mindfield® eSense Skin Response App. After that, we export the data we measured by e-mail. We will measure skin conductance with two Velcro electrodes attached to the distal phalanges of the index and middle fingers which apply a very low voltage to the fingers. Distal phalanges were chosen as the skin response amplitude is higher in this part of the finger (see Scerbo et al. 1992). We will use the dominant hand for each participant to be recorded. Although the differences in SCR seen between left and right hands are difficult to interpret, research on the topic recommend hand control experiments. As skin response is sensitive to many stimuli, the experiments will control for stress and anxiety control so as to be the only possible interpretation. Moreover, in the case of Experiments 1 and 2 we will add a few seconds between trials (slides and vignettes respectively) in order to make sure the skin response identifiable for each one of them.

Moreover, participants will be asked to respond to the Carver’s BIS/BAS theory Questionnaire (Carver & White, 1994) in order to assess their motivation levels (See Appendix II). This questionnaire measures the sensitivity to both behavioural systems (BIS and BAS), and let us know if participants move towards a reward or to avoid punishment. It is composed by 24 statements which participants should agree or disagree with. For each of them, they have to indicate accurately how much they agree or disagree with, from 1, totally agreeing, to 4, totally disagreeing. Since some studies confirm that motivating participants with a reward can produce positive effects on the quality of their performance (DePaulo et al. 1983; Vrij et al. 2000), they will be told they would win a prize if the judge is not be able to find out where they lie, and they also will be paid for their participation in the experiment. Participants will not be fully aware of the purpose of the experiment.

4.1.1. Experiment 1
In the first experiment, participants will have to choose between two pictures lying where they will be asked to do so. The cognitive load is relatively low as they did not need to tell anything else than és allà, ‘it’s there’ and pointing to an image.

4.1.1.1. Methods
4.1.1.1.1. Participants
Before starting experiment 1, a pilot experiment was run. For the pilot experiment, 5 native speakers of Catalan (1 woman, 4 men) aged between 19 -21 were recruited from the Universitat Pompeu Fabra (UPF).
4.1.1.1.2. Materials

For our pilot experiment, we used a projection of five different slides. Each slide contained two pictures of two different animals which popped up at the same time together with a one-second audio clip of the voice of one these animals. Below the images we wrote down either veritat ‘truth’ if we wanted participants to perform a truthful version, or mentida ‘lie’ if we wanted them to perform deception.

![Figure 1. Slides used in Experiment 1](image)

4.1.1.1.3. Procedure

In first experiment, participants were shown five different slides with stimulus of sounds and images of animals (see Figure 1 above). Each slide presented two images of animals together with a one-second audio clip of the sound of one the animals. The task consisted of pointing at the animal which matched with the sound if in the slide the word veritat ‘truth’ appeared in the screen, or pointing at the animal which didn’t produce the voice if the word mentida ‘lie’ appeared. While participants pointed at the animals they had to say és allà ‘it’s there’. A judge who stood in front of the participants tried to guess whether they were telling the truth or not (see Figure 2 for the experimental setup).
4.1.1.1.4. Coding and Analysis

4.1.1.1.4.1. Temporal measures of speech and gesture

The target 27 utterances were temporally analyzed. The temporal analysis of the speech and gestural data was based on the methodology by Esteve-Gibert & Prieto (2013). The gestures were divided in five different parts (see Figure 3): the *preparation phase* (movement of the arm in a relaxed position until the gestural peak of effort), the *stroke* (peak of effort in the gesture) before the apex, the *apex* (peak of maximum prominence of the gesture), the *stroke* after the apex, and the *retraction phase* (arm movement from the highest extension to the relaxed position).

![Figure 3. Illustration of the gesture phases of a pointing gesture: 1) preparation phase, 2) stroke before the apex, 3) apex, 4) stroke after the apex, 5) retraction.](image)

The stroke phase, as well as the gestural apex, was coded using ELAN (Esteve-Gibert & Prieto, 2013). The peak of the fundamental frequency and syllable boundaries of the target words were annotated with Praat. Figure 4 illustrates an example with the target tiers used in ELAN: Transcription of the speech, Gesture Phases, Apex, Pitch Peak and the coding of True and Deceptive Speech.
4.1.1.4.2. **Skin Conductance Response Measures**

Skin Response is most commonly measured by size. Size is quantified from the onset of the response to its peak and calculating the amount of increase in conductance. An elicited SCR’s size typically varies from 0.1 and 1.0 μS. (Dawson et al. 2007). Even though this measure was not taken in the pilot experiment, we will take this into account in all our studies in order to detect differences between true and deceptive utterances of the subjects.

As the stimuli are repeated, we will need to take an average size of the SCR from both the true condition and the deceptive one. There are two ways of computing SCR: Calculating amplitude or magnitude. Magnitude is the mean value of all stimuli, including those with no measurable response. Amplitude is the mean value of only the stimuli with measurable response. (Dawson et al. 2007). For our studies we used magnitude measures as it is the most frequently used and reliable. The rise time was
compared with the duration of gesture phases and pitch peaks so as to see if there is any kind of difference in alignment in truth and deception.

Dawson et al. (2007) report some problems in quantifying the skin response. There are individual differences that must be taken into account. A SCR can be considered high or low depending on each subject’s range of SCLs.

4.1.1.2. Results
We obtained a total of 27 utterances from the data, 15 from the truthful condition and 12 from the deceptive one. We calculated the average time between the deictic gestural apex and the intonation F0 peak in the two conditions. Figure 6 shows the mean temporal distance (in s.) between the gestural apex and the F0 peak in the truthful and the deceptive conditions. Results show that the average distance in time between the two locations was closer in the truth condition (0.1s) than in the deceptive condition (e.g., 0.142s. This means an increase of 29.6% in the temporal interval between the gestural apex and the F0 peak, and shows a tendency towards misalignment.

Cohen et al. (2010) found that stroke phases were shorter in deceptive speech. We also replicated this tendency. Figure 7 shows the mean duration (in s) of the stroke phase of the pointing gestures in both conditions. While the duration of the pointing stroke phase in deceptive utterances is 0.238, the stroke phase in truthful pointing is 0.26s, which represents an increase of 8.5%.

Figure 6. Mean temporal distance (in s.) between the gestural apex and the F0 peak in the truthful and the deceptive conditions.

Figure 7. Mean duration (in s.) of the stroke phase in the truthful and the deceptive conditions.
4.1.2. Experiment 2
The second experiment was conducted to elicit spontaneous deceptive stories rather than simple utterances elicited in Experiment 1. This experiment requires a higher cognitive load as the subjects will be instructed to tell actual lies.

4.1.2.1. Methods
4.1.2.1.1. Participants
Experiment 2 was performed by the same 5 UPF students (1 woman, 4 men) who previously had carried out the first experiment.

4.1.2.1.2. Materials
The second experiment was based on the study made by Cohen et al. (2010), so we used the same materials (see Figure 8). It consisted of the comic story Ivy the Terrible taken from Beano, October 19, 1996 (see Appendix I) and translated into Catalan. Participants had to read the story and then they had to tell it twice: lying and telling the truth. For that, we showed them a projection of the story without text where they had to point each vignette and tell what happened.

Figure 8. Ivy the Terrible translated into Catalan
We also prepared an instruction-sheet for the participants as they had to tell particular lies in some vignettes, as we will see in the next section.

4.1.2.1.3. Procedure

Participants were asked to read a brief comic story. They had no time limitation to assimilate the story in order to avoid cognitive load of memory effort as far as possible. For the same reason, they also had the story next to them in case they need it. The story had to be told twice: one telling the truth and one lying. In the deceptive version of the story, participants were asked to change three particular vignettes that were critical. As in Cohen et al. (2010), they were given instructions of how they had to change the three of them so as to make an attempt to reproduce the cognitive processes that takes place in premeditated deception. The changes were minor, and they were chosen so as to “incorporate and build on events depicted in the truthful version of the comic strip, as this closely emulates real-life deception.” (Cohen et al. 2010:146). In an instruction-sheet we explained the subjects the lie they had to tell, and in with vignettes, as we described below:

In Vignette 3, Ivy locks the DJ in the boot of the car. Participants were instructed to lie and tell that Ivy pushes the DJ in the back of the car, not the boot. In Vignette 7, Ivy pours liquid soap intentionally into a bubble machine so as to produce more bubbles. The subjects were asked to say that she accidentally spills the liquid soap. And in Vignette 12, many bubbles come out from Ivy’s mouth. Participants had to tell the bubbles come out of her ears.

Figure 9. From left to right, Vignette 3, 7 and 12.
As in the first experiment, they pointed at each vignette while they said, in this case, *en aquesta vinyeta* 'in this vignette' in a projection of a without-text version of the comic story, and then, they explained what happened in each vignette. Again, a judge who stood in front of the participants tried to guess which one was the false story.

### 4.1.2.1.4. Coding and Analysis

The temporal measures of speech and gesture were the same as in Experiment 1, as well as SCR measures (see section 4.1.1.1.4).

![Figure 10. Coding in ELAN.](image)

![Figure 11. On the left, participant points while he is telling a lie. On the right, participant points while he is telling the truth.](image)

### 4.1.2.1.5. Results

A total of 109 utterances were obtained, 94 from the truthful condition and 15 from the deceptive one. As we did in Experiment 1, we also calculated the average time between the deictic gestural apex and the pitch peak in both conditions. Figure 12 shows the mean temporal distance (in s.) between the gestural apex and the F0 peak in the truthful and the deceptive conditions. Results reveal an average difference of 0.042s between the two parameters. The mean time for the truthful condition is 0.21s.
between the apex and the F0 peak within the accented syllable, while the mean time for the deceptive one is 0.252. This means an increasing of a 16.7% of time that shows a tendency towards misalignment.

Figure 13 shows the mean duration (in s.) of the stroke phase in the truthful and the deceptive conditions. Again, we observe that stroke phases were slightly shorter in deceptive speech than in truthful speech (see Figure 13). Stroke phases were a mean of 0.115s shorter in deceptive utterances (a mean of 0.365s in truth and 0.25 in deception), which represents an increase of 31.5%.

4.1.3. Experiment 3
We will run within-subjects question-answer task which has not been piloted before and which deals with deception about personal information. The Guilty Knowledge Test (see Skin Conductance Response in Deception) records the skin response while participants answer different questions. Ben-Shakhar & Eitan (2003) run a meta-analysis to test the validity of the GKT with skin response measures. Results revealed many variations between studies. Several factors, like motivation or the nature of the GKT questions, influence the validity of the GKT. For that reason, we will take this information into account in order to see if the misalignment occurs in the same way when the subjects lie about personal information.

4.1.3.1. Methods
4.1.3.1.1. Materials
We will prepare two sheets of paper with a total of 10 questions in each one dealing with personal information (for example, have you got any brothers or sisters?) (See the Appendix III). The questions will be the same in both hand-outs, and participants had to answer them twice. In one of them, they have to tell the truth and in the other one, they
have to lie randomly in half of them. We will save the hand-out where participants tell
the truth as we will not need if for the task, but for the analysis so as to corroborate the
answers after finishing the experiment. For the task, the experimenter will use the
hand-out with deceptive answers and will ask the same questions to the participants,
but manipulating them in order to make the subjects answer only with “yes” or “no”. We
will also elaborate a set of ten slides with the words “yes” or “no” on each side and
alternating the positions. The slides will be projected on a screen and the subjects will
have to reply to the experimenter’s questions while they point to their answer on the
screen. The experimenter will try to guess where they lie.

In order to take SCR measures, we will use the same sensors and procedure as in the
proposed Experiments 1 and 2.

4.1.3.1.2. Procedure
We will ask the participants to answer the same 10 questions in two separate sheets of
paper. In one of them, they will be asked to write down the truth, and in the other one,
they will be asked to deceive in 5 of the answers. They will choose the questions where
they want to lie. The hand-out with the veritable answers will be taken apart as it will
not be necessary for the performance of the task. The experimenter will take the
answer sheet with deceptive responses and will ask to the participants the same
questions in a way participants only need to answer “yes” or “no” while they point at a
screen with the same answers. They will be instructed to be coherent with the answers
of the hand-out where they deceive, and tell the truth and lie in the same questions.
The experimenter will try to guess where they are lying.

4.1.3.1.3. Coding and Analysis
For this experiment the methodology will be based on Esteve-Gibert & Prieto (2013) for
speech and gestural analysis than in Experiments 1 and 2. Gestures will be divided in
ELAN in the five parts we mentioned and pitch peaks and syllable boundaries will be
annotated with Praat. For the analysis we will also take the same parameters as
reference, namely gestural apexes and peak of the fundamental frequency.

4.2. Study 2
The pilot results from the first two experiments of Study 1 have indicated that deception
may disrupt the patterns of temporal alignment between deictic gestures and speech,
together with the temporal characteristics of the stroke of deictic gestures. But, do
people perceive deception when there is a temporal disruption between for example pitch peaks and the apexes/stroke of deictic gestures? The second study will focus on the effects of different measures of gesture-speech asynchrony on the detection of deception. An experiment will be conducted to test whether the temporal misalignment of pitch peaks and deictic gestural apexes can lead to an increase in deception detection. Moreover, we would like to assess the timing patterns shown by participants in the lie detection task.

4.2.1. Methods

4.2.1.1. Participants
For this experiment, we will recruit 50 Catalan-dominant students from the Universitat Pompeu Fabra in Barcelona.

4.2.1.2. Materials
We will record 4 different audiovisual clips. Each one answers to a particular situation that will be presented in the first part of the clip. For the situations, we will elaborate four questions in Catalan that will need to be replied with four target words, namely, ahir ‘yesterday’, avui ‘today’, allà ‘there’, aquí ‘here’.
- When did you meet Mary? (Ahir ‘yesterday’)
- When did the car crash happen? (Avui ‘today’)
- Where is the meeting taking place? (Allà ‘there’)
- Where is Peter? (Aquí ‘here’)

After presenting the situation, a person will be videorecorded in a position of approximately 25-30° to the camera. The mouth of the actor will be covered so as to
participants cannot be biased by the movement of the lips. The actor will perform a word and a deictic gesture in each clip as in the following explanation:

- First clip: The actor will pronounce *ahir* while pointing backwards with the thumb.
- Second clip: The actor will pronounce *avui* while pointing to the floor with the index finger.
- Third clip: The actor will pronounce *allà* while pointing forward with the thumb.
- Fourth clip: The actor will pronounce *aquí* while pointing to the floor with the index finger.

![Figure 16](image1.jpg) Still video realization of the stroke of the pointing gesture while producing the target word *aquí*.  
![Figure 17](image2.jpg) Still video realization of the stroke of the pointing gesture while producing the target word *allà*.

After the recordings, the target videoclips will be manipulated for their audiovisual synchronization with Premiere 6.0 (Adobe Corporation) in order to create a misalignment between the deictic gestural apex and the accented syllable. For each video clip, a total of four movies will be created. One of them will display the audio and video recordings in synchrony with each other while the other three movies will display audio and video with different degrees of asynchrony. The number of frames (and ms) of separation between the audio stream and the audio stream will be decided depending on the results obtained in Study 1. A potential delay of 200ms was used in Leonard & Cummins (2011). The misalignment will be done by delaying the audio as it is not natural that the apex of the gesture occurs before the accented syllable (see Esteve-Gibert & Prieto, 2013). A potential second experiment will be run manipulating the duration of the stroke hold in the pointing gesture.

### 4.2.1.3. Procedure

For this experiment, we will use four different situations with four responses each. One of the responses will be the synchronic audio and video recordings, and the other three, the manipulations with three different degrees of misalignment done by delaying...
the audio. So, in total, we will use 16 responses: 4 synchronic movies and 12 manipulations (4 manipulated movies for each of the three degrees of asynchrony).

After showing the participants the question-situation, we will present then four responses in movie clips. One of them will be the synchronic version, and the other three, the misaligned versions. Each of the three manipulations will be delayed in a different degree that we will decide after observing the results of the Study 1. The four movies will be displayed randomly. The same situation will be shown twice. Subjects will be instructed to select between the four movies where they think there is a lie. We will not tell them there is any manipulation. This process will be repeated in each of the four situations.

5. Work Plan

<table>
<thead>
<tr>
<th>January 2017</th>
<th>Submission of the PhD. Project</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data collection for Study 1 (three experiments)</td>
</tr>
<tr>
<td>February 2017 – April 2017</td>
<td>Analysis of the data for Study 1</td>
</tr>
<tr>
<td>May 2017 – June 2017</td>
<td>Writing Study 1</td>
</tr>
<tr>
<td>July 2017-September 2017</td>
<td>Preparation of the materials for Study 2</td>
</tr>
<tr>
<td>October 2017</td>
<td>Data collection for Study 2</td>
</tr>
<tr>
<td>November 2017 – February 2018</td>
<td>Analysis of the data for Study 2</td>
</tr>
<tr>
<td>March 2018 – May 2018</td>
<td>Writing of Study 2</td>
</tr>
<tr>
<td>June 2018 – December 2018</td>
<td>Writing of the Ph.D. dissertation</td>
</tr>
<tr>
<td>February 2019</td>
<td>Dissertation defence</td>
</tr>
</tbody>
</table>

6. Selected References


DePaulo et al. (1983) investigated whether motivated liars perform better than non-motivated liars and whether they are able to control their nonverbal responses. In an experiment, two groups of participants were asked to answer four questions. In two of
them, they had to tell the truth, and in the other two, they had to lie. For the study, one of the groups was assigned to a low motivational condition and the other one to a high motivational condition. For the low motivation condition, the experiment was described to the participants as a game where they had to tell lies to the confederate and (s)he had to guess where they were lying. By contrast, for the high motivation condition, participants were told that it was really important to know how to lie and they were given an explanation about the link between the ability to lie and career success. Results revealed that high motivated deceivers performed better, as their lies were less detected.


The sympathetic nervous system controls sweat glands and response to emotional and physical changes. The Sin Conductance Response (SCR) or Electrodermal Response (EDR) is a measure of the electrical conductance of the skin which fluctuates according to stress levels. Stress raises skin conductance. The Skin response is measured by Ohm’s law principle, and conductance is measured in microSiemens (μS). In order to measure skin response, two electrodes must be placed on the phalanges of one hand. Although there is no clear proof about the effect of taking measures from dominant or non-dominant hands, it is recommended to measure conductance from the dominant one. Interestingly, the SCR is sensitive to different emotions and it is not clear to interpret measures of any particular process. However, in an experimental condition where variables can be controlled, the SCR can be interpretable. The situation of continuous stimuli of performing a task will increase the SCR.


Paul Ekman is one of the most famous experts in the field of emotions and nonverbal communication. He explains different methods and behavioural clues to detect deception. Research suggests that people acquire different types of communicative and emotional gestures. But there are some specific movements that seem to be innate reactions to emotions and they are impossible to hide. The union between these movements and verbal responses gives us the clue to assess whether emotions are coherent with people’s responses: The greater the coherence, the greater conviction, and so, the greater possibility of a truthful response. Nevertheless, when people
deceive, they may show signs of different emotions: The ones they want to hide, and the ones they want to show. This fact might trigger the appearance of discrepancy between speech and physical responses like gestures.


Cohen et al. (2010) run an experiment in order to assess the frequency of appearance of iconic gestures in truthful and deceptive speech. Participants had to read and then narrate a comic story twice. On one hand, they had to relate the story accurately, telling the truth. On the other hand, subjects were asked to change three critical details in three vignettes of the comic and they were instructed about how to change them. In that way, the experiment tries to imitate the cognitive operations deceivers do during premeditated deception. Moreover, the three changes were chosen in order to trigger the production of spontaneous iconic gestures. Results revealed that participants produced fewer iconic gestures when they were deceiving than when they were narrating the same vignettes truthfully. Results also showed that gestures during deception have fewer post stroke holds and shorter stroke phases.


Esteve-Gibert & Prieto (2013) analyzed fifteen Catalan speakers while they pointed at a screen pronouncing a set of target words with different metrical patterns in a contrastive focus condition and followed by phase boundaries. The objective was to assess 1) the alignment patterns between prominent parts of speech and gestures, and 2) whether L boundary tones affect the alignment gesture with respect to speech. Results showed that the most stable anchor for gesture apexes was the peak of the rising pitch accent. Apexes were located at the end of the accented syllable in non-phrase final position, whereas they occurred before the end of the accented syllable in phrase-final position.
General References


Appendix

I. Study 1-Experiment 2 – Ivy the Terrible (Beano, October 19, 1996)
II. Carver’s BIS/BAS theory Questionnaire:

1. A person's family is the most important thing in life.
2. Even if something bad is about to happen to me, I rarely experience fear or nervousness.
3. I go out of my way to get things I want.
4. When I'm doing well at something I love to keep at it.
5. I'm always willing to try something new if I think it will be fun.
6. How I dress is important to me.
7. When I get something I want, I feel excited and energized.
8. Criticism or scolding hurts me quite a bit.
9. When I want something I usually go all-out to get it.
10. I will often do things for no other reason than that they might be fun.
11. It's hard for me to find the time to do things such as get a haircut.
12. If I see a chance to get something I want I move on it right away.
13. I feel pretty worried or upset when I think or know somebody is angry at me.
14. When I see an opportunity for something I like I get excited right away.
15. I often act on the spur of the moment.
16. If I think something unpleasant is going to happen I usually get pretty "worked up."
17. I often wonder why people act the way they do.
18. When good things happen to me, it affects me strongly.
19. I feel worried when I think I have done poorly at something important.
20. I crave excitement and new sensations.
21. When I go after something I use a "no holds barred" approach.
22. I have very few fears compared to my friends.
23. It would excite me to win a contest.
24. I worry about making mistakes.

III. Experiment 3 - Questions:

- Have you got any brothers or sisters?
- What’s your best friend’s name?
- What is your favourite ice-cream flavour?
- Have you got any pets at home?
- What is your favourite subject?
- What would you like to do in the future? (Job)
- What is your favourite colour?
- If you could travel abroad wherever you want, where would you go?
- Do you do any sports?
- Do you prefer beach or mountain?