selected papers *of*

the **10**th **I**nternational **C**onference of **G**reek **L**inguistics

Edited by

Zoe Gavriilidou
Angeliki Efthymiou
Evangelia Thomadaki
Penelope Kambakis-Vougiouklis

DEIXIS IN INTERACTION: A MULTIMODAL APPROACH

Maria Koutsombogera

Institute for Language & Speech Processing / 'Athena' RC, Greece
University of Athens, Department of Linguistics, Greece
mkouts@ilsp.athena-innovation.gr

ABSTRACT

This paper presents a corpus-based study of deictic non-verbal expressions produced by Greek television interview participants and their relation to the co-uttered speech. Deictic non-verbal expressions were extracted from a multimodally annotated multimedia corpus and they were studied in terms of typology, modalities employed, frequency, and distribution per participant. Furthermore, in order to identify the meaningful content they acquire, they were classified into categories according to their referent, expressed or not through speech, i.e. a concrete object, person or location or an abstract concept related to time and space.

1. Introduction

In face-to-face interaction speakers employ distinct modalities in order to communicate their message. Thus, the processing of information conveyed in all modalities (gestures, facial expressions and body posture) is important both to comprehension and production of language (Ekamn & Friesen 1969, McNeill 1992, Kendon 2004). In order to study the interplay between speech and non-verbal expressions in its communicative dimension, a multimedia corpus consisting of Greek TV interviews was developed to include multimodal annotations; non-verbal modalities of hands, head, face and torso that contribute to the communication as well as the dialogue management during the interviews were identified and annotated with semiotic, semantic and pragmatic values.

Out of this rich collection of features we select to present the expressions annotated semiotically as deictic. Our goal is to (a) investigate the form and functions of deictic non-verbal expressions in an interactional context, as well as their relation with the co-uttered speech, (b) classify the non-verbal expressions according to their identified referents. For the purpose of this study we extracted all instances of gestures annotated as deictic, together with the words or phrases that are synchronized with them. Subsequently, we examine the typology of the deictic gestures, i.e. the specific forms that they may have. We then move on to the classification of the referents (e.g. self deixis, co-locutor deixis, location, time, presence or absence in the physical space, etc.) and their distribution in the data, to discover the reasons for which deixis is mostly employed. At the same time, we study the words that co-occur with the gestures to find out more about the lexical and grammatical properties that induce a deictic gesture. Finally, we examine whether the choice of the form or the trajectory of the gesture that the speakers make is determined by the respective referent category.

2. Non-verbal deixis

Deictic non-verbal expressions are mostly movements pointing to some direction with reference to concrete objects, persons or locations (present or virtual/ imaginary with regards to the physical space where an utterance takes place) or to abstract concepts related to time, space, or discourse structure. Their form does not convey semantic information itself, in contrast to other semiotic types such as the iconic ones. There are though several variations of deictic movements, i.e. in terms of the shape and the trajectory of a gesture, the direction of a gaze, etc. More importantly, they acquire meaningful content when linked to the actual referent which may or may not be expressed through speech. In this sense, they clarify the way through which the speakers refer to the surroundings.

Though deixis has been traditionally considered as the use of linguistic elements that contribute to locate entities in spatiotemporal, social and discoursal context (Lyons 1977, Marmaridou 2000,

Levinson 2006), non-verbal deixis provides access to the visible contextual information needed to understand an utterance.

Since deictic gestures, facial expressions and torso movements are a form of visible activity, their understanding and disambiguation require audio-visual information, i.e. they are dependent on both the linguistic and the extra linguistic information and context of the utterance. Non-verbal deixis has been studied within the broader context of multimodal communication and in parallel to other semiotic categories attributed to the various modalities (McNeill 1992, Kendon 2004). An efficient means to study the form and meaning of deictic non-verbal expressions is their observation through a corpus annotated with multimodal information. In this respect, non-verbal deixis has been co-investigated together with speech segments to discover their function and contribution content of the talk as well as to the turn taking system (Mondada 2007).

In principle, the classification in semiotic types follows Peirce's model of trichotomy of signs (1955) to icon, index and symbol, depending on the type of reference to the denoted object or entity. In this respect, an index is a sign that is linked to its object by an actual connection or real relation. When co-examined with language, an index is often related to a demonstrative word, though this is not always the case, as it will be explained in the following paragraphs.

3. Corpus description

The interview corpus examined is of approximately 170 minutes duration and it consists of 12 interviews. The structure is that of a typical two-person discussion program, i.e. an interviewer (host journalist) addresses questions to a guest (politicians, actors, TV personas) and the discussion follows an agenda which is predefined in either a strict or less controlled manner (Koutsombogera & Papageorgiou 2009). During the interview, the participants present their argumentation to the discussion topics and they elaborate on their opinions. In this context, they produce a multitude of non-verbal expressions co-uttered with speech.

In each interview, the audio signal has been transcribed[1] and the output has been further segmented and annotated in terms of dialogue acts[2] (Bunt et al. 2010). The video annotation[3] deals with the labeling of the non-verbal expressions (facial, hand and torso movements) co-occurring with speech, at multiple levels: (a) identification and marking on the time axis, and assignment of (b) respective semiotic type, (c) turn management type (d) feedback type, (e) attitudes and emotions expressed (W3C 2008) and (f) semantic relations with speech (Poggi & Magno Caldognetto 1996). The levels and labels used in the annotation scheme are mainly inspired by the MUMIN coding scheme notation (Allwood et al. 2007), which was created for building annotated multimodal resources to study communicative phenomena in different situations and different cultural settings, and whose categories have been proven to be reliable. Both audio and visual signals as well as the tier-subtiers annotations are perfectly synchronized; the overall set of annotation levels is distinguished by speaker, and all the annotation metadata are integrated into xml files.

The goal of this corpus building process was to study the interplay between speech and non-verbal expressions in its communicative dimension. For the purposes of the present study we focus on non-verbal expressions semiotically annotated as deictic. All instances that were annotated semiotically as *deictic* have been extracted from the corpus, and were further analyzed according to the semantic properties they express and their coordination with co-occurrent words.

4. Non-verbal instantiations of deixis in the corpus

In the TV interview corpus of 170 minutes, 10K of non-verbal expressions were annotated. Non-verbal expressions are grouped under the main modality used, i.e. hands, head/face and torso. Table 1 (left) shows the distribution of the distinct modalities in the corpus. In general, the rate of the produced non-verbal expressions is the same in all cases, independently of the interview or the participant and, on average, a speaker produces 0.9 non verbal expressions per second.

Furthermore, 3.5% of the corpus instances, i.e. approximately 350 non-verbal expressions were semiotically labeled as *deictic*. Deixis is one of the five values employed to describe the semiotic

---

[1] Transcriber (http://trans.sourceforge.net/)
[2] General-purpose functions, as defined in the draft of ISO 24617-2
[3] ELAN (http://www.lat-mpi.eu/tools/elan/)

properties of a non-verbal expression according to a Pierce-like model as included in the MUMIN coding scheme notation. The rest of the labels are related to rhythmic movements of the hands accompanying the speech (beats), to iconic and to symbolic (emblematic) expressions. A non-applicable (n/a) value was also employed for cases where there is no evident semiotic property carried by a non-verbal expression and it is attributed mostly to facial expressions and torso movements. The most frequent label attributed is the one of beats, followed by n/a, deixis, iconicity, and finally symbolic expressions, as presented in Table 1 (right).

Non-verbal expressions	%	Semiotic types	%
hands	54.5	beats	46.6
head-face	35.7	n/a	46.2
torso	9.8	**deictic**	**3.5**
		iconic	2.9
		symbolic	0.8

Table 1 Distribution of all non-verbal expressions attested
in the interview corpus (left) and their attributed semiotic types (right)

Non-verbal expressions annotated semiotically as *deictic* have been extracted from the corpus together with the words or phrases that they synchronise with, and were further analyzed in terms of their typology and the deixis referent to discover pragmatic, lexical and grammatical properties inducing a deictic non-verbal expression.

The deictic function is performed by all modalities. Specifically, 85.4% of the deictic instantiations was attributed to gestures, i.e. movement of index fingers or extending arms with palms open pointing towards a referent. Subsequently, head movements and facial expressions appear in 10.3% of the cases; gaze direction as well as head tilting combined with eyebrow raising and ranging openness of the eyes are representative expressions of a referent indexing. Finally, torso movements may also have a deictic function (4.3%) and are usually performed by turning the torso towards the referent while always combined and synchronized with another modality. While examining the distribution of deictic non-verbal expressions per speaker group, we attested that the hosts perform relatively more expressions (60.5%) than the guests (39.5%); this difference however is not statistically significant.

Deictic non-verbal expressions	%
hands	85.4
head-face	10.3
torso	4.3
host	60.5
guest	39.5

Table 2 Distribution of deictic non-verbal expressions per modality and speaker group

5. Referents of deictic non-verbal expressions

Deictic non-verbal expressions acquire meaningful content when linked to the actual referent which may or may not be expressed through speech. Furthermore, the indexing of the referent may occur with or without the use of deictic words, i.e. deictic gestures that occur in both of the following examples: *από τα δημοσιεύματα **αυτά**[4] # είναι απ'τα καλύτερα **κείμενα***. The typology of the deictic non-verbal expressions, that is, the forms they may have, were co-investigated together with the words or phrases synchronized with them. Next, we attempted to classify the referent types and their distribution in data to find out more about the reasons why deixis is employed by the speakers. We consider the deictic non-verbal expressions as egocentric, that is, the center of the deixis (origo) is the interview participant

---

[4] Examples in Greek are extracted from the corpus transcription. Words that are synchronised with the deictic non-verbal expressions are in **bold**.

who holds the floor in any given moment. Thus, the center consists of the speaker at the time and place of the utterance and the respective non-verbal expressions. The identified types are summarised in Table 3 and presented in detail in the following paragraphs.

Deixis referents	%
Self-reference	29
Other-reference	25.8
Space	9.5
Hypertextuality	8.8
Time	7.5
Discourse	19.4

Table 3 Distribution of referents of deictic non-verbal expressions

5.1 Self-reference

In general, person deixis, expressed either verbally or non-verbally, is related to the grammatical persons included in the utterance, i.e. the speaker, the addressee, the audience of the interview and the entities mentioned in the content of the utterances.

Self-reference non-verbal expressions are performed by the participants who hold the floor at any time. It is represented by 29% of the overall deictic non-verbal expressions. Speakers use their index finger or an open palm pointing at themselves whenever they want to highlight speech content that is related to themselves or their professional and personal surroundings. A pattern attested is that of stressing the first person of the personal pronoun, while at the same time the hand moves towards the speaker's chest. According to Turk (2007), speakers employ this type of gestures not only for self reference purposes, but to distinguish themselves from a group and emphasize on their own experiences as well.

Movements of this type often synchronise with forms of the 1st person of the personal or possessive pronouns: *εγώ που σας μιλώ, ως υπουργός εγώ, και με γνώρισε, εξαιτίας της δικής μου θέσης, και να με έχει ανασχηματίσει εμένα, υπέρ του υπουργείου μου, τι λέμε όμως εμείς*, etc. As Greek is a pro-drop language, the pronoun may be omitted at times; still, the speakers clearly employ deictic movements to refer to themselves: *όταν μάλιστα θεωρούμαι, ξέρω ναι, το βλέπουμε, αυτά που ζούμε, σέβομαι την αντίθετη άποψη*, etc.

Moreover, use of the second or third person is quite often linked to self-reference deictic movements to serve the goals of generalisation or expression of collectivity: *τα οποία νιώθουν και οι πολίτες* (a group where the speaker belongs), *που νομίζεις εσύ σημαντικό, να βλαβείς εσύ* (*you*, i.e. *anyone*) etc.

Besides linking of deictic non-verbal expressions to morphological forms, there are other cases of self reference. Speakers indicate explicitly or implicitly that certain events they utter are somehow connected to themselves, either as direct experiences or as part of their personal and professional surroundings. For example, while a politician is talking about his wife addressing her by her name and not revealing his relationship to her, he makes a self-reference gesture while uttering "*η κυρία Χ μένει στο σπίτι της*"; he thus implies that there exists a connection between himself and that woman.

Deictic movements are also performed while the speaker talks about his/ her own actions: *όχι κατ' εντολήν κομματικής* (orders that the speaker gives as a representative of his party), *με την προσπάθεια που γίνεται* (from a speaker who is a minister), *προσωπικά, ως ένα είδος εκτόνωσης* (for the speaker himself), etc.

5.2 Other-reference

Assuming that the speaker is the deictic center, non-verbal deixis related to persons and performed to a direction other than the body of the speaker who holds the floor are considered as other-reference. This type of reference may take two forms, deixis towards the addressee as well as towards third parties, which – in a two-party discussion – are absent from the physical space of the interview.

5.2.1 Co-locutor deixis

Non-verbal modalities employed to express deixis of the addressee (13%) are usually synchronised with uttering the co-locutor's name and/or identity, or with forms of the second person of the personal pronoun: *θα σας δώσω μια συμβουλή*, *δικαίως εσείς*, *θα σας πω εσάς άπειρο*, *σας ερωτώ εσάς*, *η εναντίον σας διαμαρτυρία*, *σας δημιούργησε πρόβλημα αυτό;* etc. Co-locutor deixis is also synchronised with larger groups or organisations to which a speaker may belong or be related to, e.g. a pointing gesture towards the minister of public order performed while a speaker utters the phrase *από την αστυνομία της Θεσσαλονίκης*.

5.2.2 Third-party deixis

Non-verbal deixis to (well-known or not) persons who are not present in the physical space of the interview, but are involved in the argumentation or mentioned in the speech content occurs quite often (11.8%): *τα πιτσιρίκια στο σχολείο*, *άλλους καλλιτέχνες*, *από το ΠΑΣΟΚ και τη ΝΔ*, *ο Ρουσόπουλος ενεπλάκη στην υπόθεση*, *εναντίον ενός υπουργού*, *αντί να το πει η Ζαχαρέα*, *η Κριστιάν Αμανπούρ*, *ο διευθυντής το κανόνιζε*, *οι φίλοι μας*, *για τις γυναίκες συναδέλφους*, *με τον πρόεδρο του κοινοβουλίου*, *αν δεν πάρουν οι δάσκαλοι*, *οι αγρότες*, *όταν ακούω από την αξιωματική αντιπολίτευση*, *ο κυβερνητικός εκπρόσωπος*, *και του δημοσιογράφου*, etc. This type of movements is assumed to be used to serve purposes of stressing the agent who carries out an action or is responsible for a statement or an event.

5.3 Space deixis

Space deixis is related to the spatial locations relevant to an utterance. As it is also the case with person deixis, the locations may be either those of the speaker and addressee or those of persons or objects being referred to.

Speakers employ deictic non-verbal expressions to indicate the whereabouts of a specific location, which rarely coincides with the interview setting (e.g. *εδώ είναι ο χώρος*, *μ' αρέσει πολύ αυτό το πλατό*, *έχουν αυτό σα στούντιο*, *είμαστε εδώ*, *το Μίμη απέξω*). In fact, most of the locations mentioned are usually related to the topic of discussion; the speakers point vaguely sideways, while uttering names of cities, districts and places related to their opinions or experiences (e.g. *η Ιρλανδία δεν πάει καλά*, *της Κρήτης στο Ηράκλειο*, *στο Νέο Δελχί*, *εγώ ήμουν στις Βρυξέλλες*, *σχολείο στο Κερατσίνι*, *στην Καλαμαριά*, *μέχρι το Βαρδάρη*, *μεταξύ Βρυξελλών και Στρασβούργου*, *στην Κρήνη*), to locations of personal and professional activities uttered by the speaker: (e.g. *μας καταγγείλατε προχτές στη Βουλή*, *το δικαστήριο από την άλλη*, *αίθουσα της ολομέλειας*), or other locations (*στο ποτάμι*).

5.4 Hypertextuality

The annotated multimodal corpus includes non-verbal expressions making reference to texts the content of which is commented upon. The term hypertextuality is adopted to describe these cases where utteramces within the discussion is linked to textual or audiovisual material within the physical space of the interview (e.g. books, pictures, reports in front of the speakers, studio displays). The deictic movements are used as a link to other texts beyond the interview text. This connection with other types of material is employed with the goal of either argumentation or information on a topic based on additional sources.

Speakers point towards these texts while uttering words such as: *αυτό που είναι συγκινητικό είναι*, *εδώ γράφει αγγλική φιλολογία*, *είναι ένα από τα καλύτερα κείμενα*, *υπάρχουν και αντιφάσεις μέσα στις καταθέσεις*, *αυτή η συνάντηση* (an event presented within the reports that the speaker has in front of him), *στον ίδιο ακριβώς χρόνο*, *βλέπω κάτι φωτογραφίες εδώ*, etc.

Moreover, speakers often point to displays located in the studio where the interview takes place, e.g. in cases of broadcast commentaries or event coverage related to the topic of the discussion: *έτσι λέτε*, *το 'πατε*, *θα επιμείνω στη δήλωση*, *όπως είπε και ο Χ*, *θέλω να τις δούμε*, *όπως είπε και ο συνδικαλιστής*, *από όλα αυτά τα οποία παίξαμε*, *έλεγε η τηλεόραση*, etc.

5.5 Temporal deixis

Non-verbal expressions related to temporal deixis are linked to aspects of time involved in the events included in the utterances. In this sense, deictic movements of temporal reference attested in the corpus are related mainly to the past. The trajectory and direction of the movements in question are usually backwards or sideways: *χτες κάναμε μια εκπομπή* (head, hands and torso move sideways), *όσο **στην πρώτη τετραετία*** (both hands move sideways), *το είπε ο κύριος **πριν**, μιας Κυριακής **πριν από αρκετά χρόνια**, γι' αυτό **έβαλαν τη σύζυγό μου*** (hand moves backwards), *το Δεκέμβρη τώρα* (hand moves backwards), etc. Reference to the present is more rare, e.g.: *και **τώρα*** (hand moves down).

5.6 Discourse deixis

Discourse deixis is used to indicate in an abstract manner the introduction of new topics in the interview. Specifically, new topics are considered as metaphoric images or locations to which the deictic movements point; they are employed by the speakers with the goal of presenting the co-locutors with a new material as a possible forthcoming topic of discussion, e.g. ***πάμε τώρα να δούμε**.*

Unlike its function with regards to the verbal content, non-verbal discourse deixis does not refer to parts of the discourse; if a speaker wishes to make reference to an utterance or a topic that has been previously mentioned, then the movements used are related to the time slot when the utterance or topic took place (i.e. the past) and not the topic itself, e.g. in the utterance *είπατε ότι* [...] the hand moves slightly backwards to make a temporal reference.

6. Patterns in deictic movements

Although there are variations in the intensity and form of deictic movements, it is possible to group together certain non-verbal expressions in the sense that their direction, when linked to specific speech content, could lead to the formulation of patterns. The most prominent example is that of the use of the index finger, the thumb or the whole palm pointing or moving backwards when referring to events or participants of events of the past.

Moreover, deictic movements to sideways indicate referents that are far from the speaker (e.g. either locations or persons not in the physical interview space). However, there was no evidence to support that the number of hands (single vs. both) used to perform this kind of gestures entails specific correlations to the speech equivalent. Finally, speakers often point down when accompanying a specific demonstrative word (e.g. *και **αυτό** με ρωτούσε, **εκείνη** η συνθήκη, **αυτόν** τον ξυλοδαρμό*) or adjectives and adverbs of spatiotemporal properties: *και **τώρα**, από το **σημερινό** πλαίσιο, είμαστε **εδώ**, **τελευταίο** επεισόδιο*.

7. Conclusions

The corpus-based study of deictic non-verbal expressions revealed that there are multiple ways through which deixis is performed, as well as multiple types of referents to which the deixis is addressed. A major contribution of the study of non-verbal deixis is its employment by speakers when there is no actual demonstrative word or what is considered as a traditional deictic expression, be it a temporal, spatial or discourse vocabulary. In fact, we believe that in order to come up with an as complete as possible framework for deixis interpretation, both verbal and non-verbal instantiations should be studied in their inter-relation and complementarity.

Furthermore, when it comes to communication, deixis can be a powerful means to interpret the dynamics of a given interaction. Specifically, both verbal and non-verbal signs deictic signs can account for social goals of the speakers; kinesic features of a deictic movement such as the proximity of a gesture to/ from a speaker or its direction, when expressed in parallel to the speech content, can be considered as devices that both unveil a speaker's intention and enable him/ her to achieve a social goal.

Our future goal is to provide elaborate evidence towards the development of an interpretation framework of deixis. Specifically, we plan to enrich our data with broader communicative instances of face-to face interaction in order to strengthen our findings regarding forms and variations of non-verbal deixis, as well as to further elaborate on the established categories of referents related to deictic movements.

References

Allwood, J., Cerrato, L., Jokinen, K., Navarretta, C. and Paggio, P. 2007. "The MUMIN Coding Scheme for the Annotation of Feedback, Turn Management and Sequencing Phenomena" *Journal on Language Resources and Evaluation* 41(3–4):273–287.

Bella, Spyridoula. 2005. "Cognitive motivation and pragmatic functions of the Greek deictics." *Journal of Greek Linguistics* 6(1):39–60.

Bunt, Harry, Alexandersson, Jan, Carletta, Jean, Choe, Jae-Woong, Fang, Alex C., Hasida, Koiti, Lee, Kiyong, Petukhova, Volha, Popescu-Belis, Andrei, Romary, Laurent, Soria, Claudia, and Traum, David R. 2010. "Towards an ISO Standard for Dialogue Act Annotation." In *Proceedings of LREC 2010- Language Resources and Evaluation Conference*, Malta.

Ekman, Paul, and Friesen, Wallace V. 1969. "The Repertoire of Nonverbal Behaviour: Categories, Origins,Usage and Coding." *Semiotica* 1:49–98.

Kendon, Adam. 2004. *Gesture: Visible Action as Utterance*. New York: Cambridge University Press.

Koutsombogera, Maria, and Papageorgiou, Harris. 2009. "Multimodality Issues in Conversation Analysis of Greek TV Interviews." in *Multimodal Signals: Cognitive and Algorithmic Issues*, edited by Anna Esposito, Amir Hussain, Maria Marinaro, and Roberto Martone, LNAI 5398:40–46. Berlin/ Heidelberg: Springer.

Marmaridou, Sophia. 2000. *Pragmatic Meaning and Cognition*. Amsterdam & Philadelphia: John Benjamins.

McNeill, David. 1992. *Hand and Mind: What Gestures Reveal About Thought*. Chicago: University of Chicago Press.

Mondada, Lorenza. 2007. "Multimodal resources for turn-taking. Pointing and the emergence of possible next speakers." *Discourse Studies* 9(2): 194–225.

Levinson, Stephen C. 2006. "Deixis." In *The Handbook of Pragmatics*, edited by Laurence R. Horn and Gregory L. Ward, 978–120. Blackwell Publishing.

Lyons, John. 1977. "Deixis, space and time." *Semantics* 2: 636–724.

Peirce, Charles S. 1955. "Logic as Semiotic: the theory of signs." In *Philosophical Writings of Peirce*, edited by Justus Buchler, 98-115. New York: Dover Publications.

Poggi, Isabella, and Magno Caldognetto, Emanuela. 1996. "A Score for the Analysis of Gestures in Multimodal Communication." In *Proceedings of the Workshop on the Integration of Gesture and Language in Speech, Applied Science and Engineering Laboratories*, edited by L.S. Messing, 235–244, Delaware: Newark and Wilmington.

Turk, Monica. J. 2007. "Self-referential gestures in conversation." *Discourse Studies* 9(4): 558–566.

W3C Incubator Group. 2008. "Emotion Markup Language – Report." Accessed February 14, 2012. http://www.w3.org/2005/Incubator/emotion/XGR-requirements-20080513/.