UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

The Middle Voice in Ancient Greek. A study in Polysemy

Allan, R.J.

Publication date
2002

Link to publication

Citation for published version (APA):
Allan, R. J. (2002). *The Middle Voice in Ancient Greek. A study in Polysemy*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:17 Nov 2026

CHAPTER 1

Introduction

1.0 The Problem

The puzzling diversity of the different usage types of the middle and passive voice[1] in Greek has been the subject of debate ever since the times of the ancient grammarians. The nature of the problem becomes apparent when considering the brief inventory of middle uses given by Rijksbaron (cf. 1994[2]: 159-60, slightly adjusted).[2]

(i) **Passive use**
The patient is the subject. The aorist is of the passive type (i.e. in -(θ)η-).

(1) αἱ νῆες παρεσκευάσθησαν ὑπὸ τῶν Ἀθηναίων
'The ships were prepared by the Athenians'

(ii) **Direct reflexive use**
This use occurs with agentive transitive verbs, typically verbs of grooming. The subject-referent performs the action on himself. Agent and patient are co-referential. In contrastive contexts a reflexive pronoun is added as direct object. The aorist is of the sigmatic middle type. E.g.:

(2) a. ἐλουσάμην (ἐμαυτόν) 'I bathed (myself)'
b. παρεσκευασάμην (ἐμαυτόν) 'I prepared myself'

(iii) **Indirect reflexive use**
This use occurs with agentive transitive verbs. The subject-referent performs the action in his own interest. The aorist is of the sigmatic middle type. E.g.:

(3) παρεσκευασάμην τὰς ναῦς
'I prepared the ships in my own interest'

---

[1] With the term *middle voice* I refer to the *inflectional* category in Greek, i.e. in the present stem the middle voice includes middle-inflected verbs with passive meaning. The passive voice (formed with the suffix -(θ)η-) is only found in the aorist and future stem.

[2] This inventory is only meant as a brief survey of the different middle uses. In ch. 2, I will describe the middle uses more extensively.

(iv) **Pseudo-reflexive and pseudo-passive use**
The corresponding active transitive verbs are causative. The middle verb denotes either that the subject brings about a change of state[3] to himself (pseudo-reflexive; typically verbs of movement), or that the subject undergoes a change, no agent being present (pseudo-passive; typically verbs of emotion/cognition and physical processes). The aorist is of the passive type. E.g.:

(4) a. ἀπηλλάχθην 'I went away'
b. ἐφοβήθην 'I became afraid'
c. ἐτάκην 'I melted'

(v) **Media tantum and Passiva tantum**[4]
The middle only verbs have an aorist of the middle form, and are agentive. The passive only verbs have an aorist of the passive form, and are mostly non-volitional. If so, they are similar to *pseudo-passives*.

(5) *Media tantum*: ἠτιασάμην αὐτόν 'I accused him', ἠλάμην 'I jumped'
Passiva tantum: ἥσθην 'I enjoyed myself'

Similar enumerations of middle uses are found in all grammars of Ancient Greek, e.g. Goodwin (1895: 265-8), Kühner-Gerth (I: 100-29), Gildersleeve (1900-11: 64-79), Stahl (1907: 42-74), Smyth-Messing (1956: 389-398), Schwyzer-Debrunner (II: 228-41), García Gual (1970), and Martínez Vázquez, Ruiz Yamuza & Fernández Garrido (1999: 229-253).[5] However, such lists almost always leave the reader with the following questions:[6]

---

[3] The notion *change of state* implies that an entity is in a certain state at moment T_1 that is different from the state it was in at T_0. This means that, for instance, a verb like *hit* does not denote a change of state, since the direct object may be in the exact same state after it was beaten as it was before. Verbs that do denote changes of state are *break*, *melt*, *destroy*, *remove*, *frighten*, *persuade*. Changes of state can be subdivided into physical and mental changes of state, and changes of location.

[4] Media tantum (or *middle-only* verbs) are middle verbs that do not have active counterparts. They are sometimes called *deponentia*, a term borrowed from Latin grammar. This term is less adequate since it suggests that these verbs have "laid off" (i.e. lost) their active forms. There is no historical evidence that this is what actually happened. In the present study, the term *media tantum* (or *middle-only* verbs) will be used as a covering term to refer both to *media tantum* in the strict sense (i.e., verbs with middle aorist forms) and to *passiva tantum* (verbs with passive aorist forms), cf. Rijksbaron (1994[2]: 157). I am aware, however, that this morphological dichotomy has important implications, especially regarding the semantics of the verbs concerned. In fact, chapter 3 will be devoted almost exclusively to the relation between form and meaning in the aorist stem.

[5] Other contributions to the study of the middle (and passive) voice in Ancient Greek are: Carlsson (1912), Chantraine (1927), Marguliés (1929-30), Wistrand (1941), Vaillant (1942-5), Schwyzer (1943), Vendryes (1948), Blass-Debrunner-Rehkopf (1979[15]: 255-65), Hartmann (1954), Humbert (1954[2]: 103-9), Boeder (1961), Bechert (1964), Zsilka (1964), Zsilka (1965), Benveniste (1966), Jahnkuhn (1969), Bader (1972), Plénat (1974), Barber (1975), Michelini (1976), Cock (1981), Ruipérez (1986), Risselada (1987), Delaunois (1988: 79-93), Anderson (1993), Bakker (1994), Perel'muter (1995), Sicking & Stork (1996: 131-7), Basile (1998: 320-32), Kurzová (1999).

(I) **Is there a semantic element common to these usage types? If so, how should it be defined?**

(II) **Assuming that the various middle uses constitute a polysemous structure: in what way are the middle uses related to one another?**

(III) **What is the explanation of the differences in aorist formation?**

Until recently, the middle voice has been the exclusive domain of classical and Indo-European linguists. During the last decades, however, the middle voice has become the object of increased interest in typological research. One of the factors that brought about this increased interest is the striking similarity that middle voice systems display across languages, even in those which are genetically non-related. Of great importance to the recent fruitful research on middle voice phenomena is the insight that *reflexive systems* - as those found in many of the modern European languages - are functionally highly similar to (inflectional) middle voice systems as we find them in Ancient and Modern Greek and Sanskrit. On account of this remarkable functional similarity between reflexive and middle voice systems (here used in the strict sense), Kemmer (1993) subsumes both grammatical systems under a single, semantically defined notion *middle voice*. Middle voice systems and reflexive systems are not only found in ancient languages like Greek, Latin and Sanskrit, but also in modern European languages like, for example, Romance, Slavic, Germanic (English being a notable exception), and in non-Indo-European languages, as in Tamil (Dravidian) and Fula (Niger-Congo), see Klaiman (1991).

Major contributions to our knowledge of middle voice phenomena, most of which are from a cross-linguistic standpoint, are: Faltz (1977), Shibatani (1985, 1988), Lichtenberk (1985), Geniušiene (1987), Croft, Shyldkrot & Kemmer (1987), Klaiman (1988, 1991, 1992), Kemmer (1993, 1994), Fox & Hopper (1994), and Givón (1994). Besides these general works on middle voice phenomena, several important language-specific studies have appeared during the last decades. These include the works of Gerritsen (1990, 1992) [Russian], Maldonado (1999) [Spanish], and Manney (2000) [Modern Greek].[7]

Apart from studies that are concerned exclusively with middle-reflexive systems, there have been a number of important developments in linguistic theory that may increase our insight into the middle voice as a grammatical category. These developments will be the topic of the following sections. In section 1.1, I will discuss the notion of Prototypical Transitivity (Hopper & Thompson 1980) and its crucial importance to voice distinctions. In section 1.2, the nature of complex polysemous categories (Lakoff 1987, Langacker 1987, 1991a, 1991b, 2000) and the notion of category prototype will be discussed.

---

[6] Gonda (1960: 31) expressed his discontent about the existing treatments of the middle voice in the following manner: "It is true that the old-fashioned enumerations of the functions of this category in class-books (...) created the impression of an incoherent mixture of functions, and that attempts to give a general definition laboured under delusions and vagueness: (...)".

[7] Comprehensive studies on the semantics of other ancient Indo-European languages are: Neu (1968a, 1968b) [Hittite], Flobert (1975) [Latin], Gonda (1979) [Vedic Sanskrit], Schmidt 1969 [Tocharian].

1.1 Voice: some basic notions

1.1.1 The Prototypical Transitive Clause

Voice alternations can be characterized as encodings of different choices of clausal subject (cf. Langacker 1991a: 335). For example, the familiar passive construction can be described as a marked expression of the special configuration in which the expected subject is bypassed in favour of a less typical subject. It has often been observed that voice alternations can be fruitfully described as markings of departures from the *prototypical transitive event* (Hopper & Thompson 1980; Givón 1984: 157, 2001a: 126-8; Langacker 1991a: 335).[8] The importance of the notion of prototypical transitivity for grammatical organization was first fully recognized by Hopper and Thompson in their seminal article published in 1980. On the basis of vast typological evidence they claimed that the grammatical category *transitive* is structured around a prototype that can be defined by means of a cluster of semantic properties. These properties are listed as follows (Hopper & Thompson 1980: 252: the terminology used is theirs):

Table 1: *Semantic properties of the prototypical transitive clause*

	HIGH Transitivity	LOW Transitivity
A. Participants	2 or more participants	1 participant
B. Kinesis	action	non-action
C. Aspect	telic	atelic
D. Punctuality	punctual	non-punctual
E. Volitionality	volitional	non-volitional
F. Affirmation	affirmative	negative
G. Mode	realis	irrealis
H. Agency	A high in potency[9]	A low in potency
I. Affectedness of O	O totally affected	O not affected
J. Individuation of O	O highly individuated	O non-individuated

Clauses can be ranked on a scale of transitivity according to whether they have properties in common with the Prototypical Transitive Clause. In other words, clauses can be *more* or *less* transitive. Hopper & Thompson (1980: 253) give the following examples :

(6) a. *Jerry likes beer*
b. *Jerry knocked Sam down*

[8] In this study, the term 'event' is used in a broad sense subsuming several, more specific, event types such as *states* (static; non-volitional), *processes* (dynamic; non-volitional), and *actions* (volitional). Thus it is roughly equivalent to Dik's term *State of Affairs* (Dik 1987).

[9] Following Dixon's example (Dixon 1979), A and O denote Agent and Object. These notions roughly coincide with transitive agent and patient.

Example (b) is much higher in transitivity than (a) because it displays the following properties:

- Kinesis: *action*
- Aspect: *telic*
- Punctuality: *punctual*
- Affectedness of O: *total*
- Individuation of O: *high; referential, animate, and proper.*

For each of the semantic properties of the prototypical transitive clause (Table 1), Hopper and Thompson have cited languages in which clauses that possess this particular property display transitive behaviour. Conversely, clauses that differ from the Transitive Prototype with regard to only one of these properties are found to display *intransitive* behaviour, even if they possess other prototypical transitive properties.

Of the features enumerated by Hopper and Thompson, Givón (1984: 20) considers two of primary importance: *agency* and *affectedness*. In this way, the prototypical transitive is primarily conditioned by:

(7) a. the presence of a visible, volitional, and controlling **cause/agent**; and
b. the presence of a clearly visible result-registering **effect/patient**.

Furthermore, according to Givón, agency and affectedness (and, therefore transitivity) are strongly conditioned by the clausal property of *perfectivity*. The connection between these notions can be described as follows (Givón 1984: 157; italics are his):

(8) a. *Affectedness of patient*: "The more completed an event is, the more likely it is that the patient in fact *registers* to the full the effect of the action".
b. *Effectiveness of agent*: "The more successfully completed the event is, the more likely it is that the agent *was* in fact the deliberate, direct, effective cause of that successful completion".

Examples of prototypical transitives given by Givón (1984: 20) are:[10]

(9) a. *Mary cut the meat*
b. *John destroyed the house*

1.1.2 Langacker's Billiard-ball Model

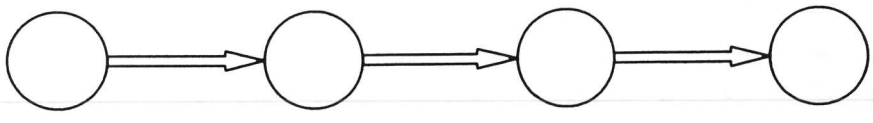
The notion of prototypical transitivity also plays a major role in Langacker's *Billiard-ball model*. The Billiard-ball Model is a *cognitive model*. In Cognitive Grammar, it is claimed that meanings are to be analyzed in relation to *cognitive domains*. The notion of cognitive domain can be defined as "[a] context for the characterization of a semantic unit" (Langacker 1987: 147). For example, the meaning of the word *glass* evokes the cognitive domains of *space* [in which all concrete objects exist], of *shape* [typically cylindrical],

[10] More examples can be found in Givón (2001a: 127).

material [glass], *size* [easily held in the hand], but also more complex domains such as its typical *function*. The typical function of a glass could be characterized as an instrument used in the process of drinking: it is filled with a liquid that is suitable for human consumption, it is grasped and lifted with the arm, etc. Cognitive domains can, therefore, be viewed as knowledge bases relative to which the meaning of an expression can be assessed.[11] The notion of cognitive domain is very similar to Lakoff's notion of Idealized Cognitive Model (Lakoff 1987), and to other notions such as *frame* (Fillmore), *scene*, *schema*, or *script*.

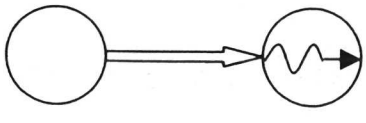
The character of a cognitive domain can be of a more idealized or archetypal nature. Such an idealized cognitive domain can be referred to as a cognitive *model*, which brings us back to the Langacker's Billiard-ball Model. The Billiard-ball model is an archetypal cognitive model that structures our conception of events. The model conceives the world as containing discrete objects which are constantly moving around, making contact with one another, and participating in energetic interactions (Langacker 1991a: 283). As a result of the interactions of different objects an *action chain* arises:

Fig. 1 *Action Chain* (Langacker 1991: 283)



One entity (represented by a circle) is charged with energy, and makes forceful contact with a second entity. As a result, energy is transmitted from the first source-entity to the second (this is shown as a double arrow). This entity is thereby driven into contact with a third, which again brings about a transmission of energy. This process can continue indefinitely, until, finally, an entity absorbs the energy and thereby undergoes a change of state. The simplest instance of an action chain is one in which the initial energy-source (which Langacker calls the *head* of the action chain) interacts directly with the final energy-sink (the *tail* of the action chain). This configuration is depicted in Figure 2:

Fig. 2 *The Prototypical Transitive Event*



Agent Patient

[11] Typically, as we have seen in the case of the word *glass*, a linguistic unit invokes multiple cognitive domains, and the number of domains evoked cannot be sharply delimited. This view can be characterized as the *encyclopedic* view of semantics (as opposed to the *dictionary* view). In the encyclopedic view, it is claimed that there is no specific boundary between linguistic and non-linguistic knowledge (cf. Langacker 1987: 154-166).

Here, the action chain head is the agent, and its tail the patient. In the Prototypical Transitive Event, the agent and the patient figure in their most prototypical role. Langacker characterizes agent and patient as follows:

(10) "The archetypal **agent** is a person who volitionally initiates physical activity resulting, through physical contact, in the transfer of energy to an external object. Its polar opposite is an archetypal **patient**, an inanimate object that absorbs the energy transmitted via externally initiated physical contact and thereby undergoes an internal change of state" (Langacker 1991a: 285)

The patient's change of state is depicted by a squiggly arrow in Fig. 2. It should be noted that Langacker's charactization of agent and patient concerns *archetypes*. Obviously, not every clause conforms to these archetypes since human experience is too rich and varied. For instance the archetypal patient, as described by Langacker, concerns an object that is *affected* by the event. However, an important class of patients that are similar to this archetype concerns *effected* objects. These are objects that have no prior existence but rather are created by the event itself, as in *She built a house*. Effected objects are generally not distinguished from affected objects by any special marking (see Langacker 1991a: 362).

The concept of the prototypical transitive event is also highly relevant to the grammatical relations of *subject* and *object*. The unmarked coding of the prototypical event is that the subject is the agent, and the object is the patient (see Langacker 2000: 25, Givón 2001a: 126, *inter al.*). Other, less-prominent participants in the event, such as instruments or experiencers, are typically coded by oblique cases or prepositional phrases.

This brings us to the issue of semantic roles.[12] There is much controversy on the definitive number of semantic roles and their exact definition. Every linguistic theory, or perhaps even every individual linguist will posit a different inventory of roles that are, furthermore, defined differently. However, these inventories generally display a considerable similarity. In this study, Langacker's description of a number of archetypal semantic roles is adopted.

The list of semantic roles enumerated by Langacker (1991a: 285-7) is as follows:

(11)
- *agent* and *patient* (see definition above)
- *instrument*: a physical object manipulated by the agent to affect the patient, thereby serving as an intermediary in the transmission of energy
- *experiencer* an animate entity engaged in a mental event
- *zero*: an entity that merely occupies a location or exhibits a static property.

Langacker's role archetypes listed above are of a highly abstract character.[13] This is a reflection of their cognitive fundamentality and their ubiquitous manifestation in the

[12] Semantic roles are also known as *case roles*, *thematic relations*, *theta-roles*, *notional roles*, and *semantic functions*.

[13] Langacker also distinguishes the semantic role of *mover*. The mover, according to Langacker, is an entity that undergoes a change of location, as *Rocky* in *Rocky drove to the beach* and *the door* in *I opened the door*. In this study, however, moving participants will be viewed either as *agents* (if the motion is volitional) or as *patients*.

languages of the world. However, Langacker (1991a: 285) admits that these roles do not form an exclusive club, and that finer distinctions can be made. A number of additional roles are useful in an adequate description of the semantics of the middle voice in Greek. These roles, I suspect, will not be very controversial:

(12) • *beneficiary*: an animate entity receiving benefit as a result of the event.
• *recipient*: an animate entity into whose possession something is transferred.
• *cause*: an inanimate entity that causes a physical or mental change in another entity.
• *source*: the location from which an entity moves, or, metaphorically, the stimulus of a mental process.
• *goal*: the location towards which an entity moves.

It should be noted that many accounts of semantic roles do not distinguish the cause-role from the instrument role (cf. Givón 1984: 126, 2001a: 161-2; Palmer 1994: 5). A cause is similar to Dik's semantic function *Force* (Dik 1997a: 118; cf. also Van Valin & LaPolla 1997: 85). The cause-role differs from an instrument in that it does not imply the presence of an initiating agent: the cause itself is conceptualized as an autonomous initiator. In Greek the cause is typically expressed by the dative case (*dative of cause)*, but a cause can also have the status of subject in a transitive clause. Since causes are autonomous initiators (i.e. only differing from agents in volition), this is not surprising. For instance,

(13) a. ὁ σίδηρος τῷ πυρὶ$_{\text{CAUSE}}$ τήκεται 'The iron is being melted by the fire'
b. τὸ πῦρ$_{\text{CAUSE}}$ τήκει τὸν σίδηρον 'The fire is melting the iron'

The cause-participant may appear in mental events, e.g.,

(14) ἥσθην ἀπειλαῖς$_{\text{CAUSE}}$ (Ar. *Eq.* 696)
I was delighted at your threats

A source is basically a spatial role, but it can also appear in an extended metaphorical sense. Examples of mental sources (or *stimuli*) are the genitival complements of verbs of perception, emotion, and cognition:

(15) a. *Perception*: ἀκούσαντες τῆς σάλπιγγος$_{\text{SOURCE}}$ (Xen. *An.* 4.2.8)
Having heard the trumpet (...).
b. *Emotion*: ὁ ἀνὴρ (...) κήδεται τῶν Θηβαίων$_{\text{SOURCE}}$ (Xen. *HG* 6.4.5)
The man (...) cares about the Thebans.
c. *Cognition*[14]: μάθε δέ μου$_{\text{SOURCE}}$ καὶ τάδε (Xen. *Cyr.* 1.6.44)
Learn from me this too.

In the above, the notion of prototypical transitive event, which involves a physical transmission of energy from an agent to a patient, was set out. It is, however, a widely-

[14] With verbs of cognition and perception, the genitival complement can also denote the entity *about which* one perceives or learns something, e.g. πατρὸς ἀκούσας (δ 114) ('having heard about father'). This type of genitive is often viewed as a partitive genitive.

occurring phenomenon that the transitive clause structure is extended to code other situation types. An example of this phenomenon is the way mental events (perception, cognition, and emotion) are treated. Examples of these are 'see', 'know', 'understand', 'want', and 'love' (see Langacker 1991a: 303-4, 310, 2000: 26). These events are coded in English as a transitive with a subject and a direct object, although they obviously do not involve a physical transmission of energy.

(16) I {see/know/understand/want/love} it (Langacker 1991a: 303)

This extension, from the prototypical transitive event to the mental event, has a metaphorical character. Its motivation can be found in the abstract commonality that is inherent in both types of events. On the one hand, we have the transmission of energy from an active initiator (the agent) to a passive endpoint (the patient), and on the other hand, we have the concept of a metaphorical *mental* path leading from a more active, conscious participant (an *experiencer*) to a more passive object-participant. In other words, mental phenomena such as gazes and direct attention can be conceived of as paths, analogical to a physical path like that of an energy flow.[15] Examples in Greek of mental events coded as transitive constructions are numerous: γιγνώσκω τι 'I realize sth.', οἶδά τι 'I know sth.', ὁράω τι 'I see sth.', τρέω τι 'I fear sth.'.[16]

1.1.3 Prototypical Transitivity and the Middle Voice

In the previous section it was observed that in the unmarked case, the agent of the prototypical transitive is coded by the subject, and the patient by the direct object. Consider now the following clauses (from Langacker 1991a: 335):

(17) a. *He opened the door.*
b. *The door opened very easily.*
c. *The door suddenly opened.*
d. *The door was opened.*

Here we can observe that in the unmarked prototypical transitive clause (a), the agent is subject, the patient is object, and the verb is in the active voice. The other clauses depart from this prototype in one crucial respect: the *patient* is coded as the subject. This departure resides in the fact that the participant that is expected to be the subject (the agent) is bypassed in favour of a less qualified candidate: the patient. Clauses (b) and (c) are formally in the active voice, but they designate only the patient's participation. In (b) the adverb *very easily* implies the efforts of an unspecified agent, which are facilitated by the inherent characteristics of the patient-subject (the door). In (c), the implicit reference to an agent is non-salient and may be absent altogether. Clauses like (b) and (c) are

[15] Often, verbs of perception and cognition are based on a metaphor, e.g. English *perceive* from Latin *-cipiō* 'grasp'; English *grasp* 'understand'; Dutch *be-grijpen* from *grijpen* 'grasp'. On the importance of metaphor to the verbs of perception and cognition, see Sweetser (1990).

[16] The non-prototypicality of these transitive verbs is manifested by the fact that they do not have a passive (οἶδα, τρέω), or acquire a passive only after Homer's time (γιγνώσκω, ὁράω).

sometimes called cases of middle *diathesis*, since they are formally in the active voice (at least in English), but *semantically* belong to the middle domain. In clause (d), though the agent is left unspecified (as is the case in most passive clauses), the efforts of an agent are definitely implied. In this respect the passive clause (d) differs from clauses (b) and (c). In the case of the English passive the deviation from the prototypical transitive arrangement is marked by the alternation of voice: instead of the unmarked active voice the verb has the marked passive voice.[17] In the example of the familiar English passive, it can be observed that the notion of prototypical transitivity is highly relevant to voice alternations. Below, I will argue that the prototypical transitive event is also essential to the characterization of the middle voice in Ancient Greek (cf. Question (I) in (1.0)).

There have been many attempts to capture the essence of the semantics of the Greek middle voice. This is not an easy task if one considers the diversity of middle usages such as passive, intransitive[18], direct reflexive and indirect reflexive. It is clear that a core-meaning, if there is any, could only be of a highly abstract nature. It is, therefore, useful to consider some of the most important definitions of the middle voice.

The natural starting-point of this overview is, of course, the unrivalled grammar of Kühner and Gerth (K-G):

(18) "Die Medialform bezeichnet eine Thätigkeitsäusserung, welche von dem Subjekte ausgeht und auf dasselbe wieder zurückgeht. Diese von dem Subjekte ausgehende und auf dasselbe wieder zurückgehende Thätigkeitsäusserung kann entweder bloss auf das Subjekt beschränkt sein, als: βουλεύομαι, ich berate mich, λούομαι, ich wasche mich, oder auf ein Objekt seiner Sphäre, (...), als ἐκοψάμην τὴν κεφαλήν ich schlug mir das Haupt, κατεστρεψάμην τὴν γῆν, ich unterwarf mir das Land (...)." (K-G, I: 100)

K-G's definition is of a remarkable originality since, as we shall see, it differs strongly from later definitions. However, it is interesting to note that K-G's definition seems to convey an imagery that is related to modern views in which the *middle voice* is claimed to express that the subject is the *Startingpoint/Initiator* (cf. "ausgeht") as well as the *Endpoint* ("zurückgeht") of the action chain (cf. Croft, Shyldkrot & Kemmer 1987; for similar views: Klaiman 1988: 26-27, Klaiman 1991, Gerritsen 1990: 11, Croft 1994: 103). I will elaborate on this matter later. However, as is the case with these more recent views, K-G's characterization only concerns the middle in the strict sense, i.e. it does not include the passive middle.[19]

---

[17] The main function of the passive construction is often taken to be to defocus the agent (see Shibatani 1985), or, as Givón puts it: "The agent is extremely non-topical ('suppressed', 'demoted'), so that the patient is the surviving topical argument in the clause" (Givón 2001b: 94).

[18] Here used in the semantic sense. The intransitive middle covers Rijksbaron's pseudo-reflexive and pseudo-passive.

[19] In modern linguistic literature there is much confusion about the terms *voice* and *diathesis*. Many authors seem to use the terms interchangeably. Others seem to use *diathesis* as pertaining to the formal properties of the verb (e.g. Klaiman 1988: 27; Duhoux 2000: 117), and *voice* as pertaining to the semantic properties of the verb. Again, other authors use the terms the other way around (e.g. Geniušiene 1987). In

More familiar formulations of the meaning of the middle voice are:

(19) "The Middle Voice denotes that the subject is in some especial manner involved or interested in the action of the verb." (Gildersleeve 1900: 64)

(20) "Verba (...), die ihren Schauplatz in der Sphäre des Subjekts haben, bei denen das ganze Subjekt als beteiligt scheint." (Brugmann 1903: 104)

(21) "Dans l'actif les verbes dénotent un procès qui s'accomplit à partir du sujet et hors de lui; dans le moyen, qui est la diathèse à définir par opposition, le verbe indique un procès dont le sujet est le siège; le sujet est intérieur au procès." (Benveniste 1966: 172)

(22) "En indo-iranien et en grec les désinences moyennes indiquent que le sujet est intéressé d'une manière personelle au procès." (Meillet 1937: 244)

(23) "The implications of the middle (when it is in opposition with the active) are that the 'action' or 'state' affects the subject of the verb or his interests." (Lyons 1969: 373)

With regard to these definitions one can object either that they seem to focus on the indirect reflexive meaning of the middle (Gildersleeve and Meillet: the subject's interest), or, conversely, that they seem to be more suitable for the passive and intransitive uses of the middle (Brugmann and Benveniste: "the whole subject participates" and "the subject is internal to the process", respectively). Only Lyons' characterization pertains clearly to both the 'passive' meanings ("affects the subject") and the indirect reflexive meaning ("or his interests"). Lyons' recognition of these two extremities of the meaning of the middle, and the transparency of his formulation, make his definition appealing. Lyons uses the verb *affect* in a sense that is broad enough to be applicable to all middle meanings. It is this broad sense of *affect* and *affectedness* that we need in an adequate definition of middle meaning.[20] If we are tempted to interpret *affect* and *affectedness* in a narrower sense, that is, as an equivalent of the ancient term πάθος, we inevitably run into trouble, since the

---

the Latin grammatical tradition *vōx* (the ancestor of *voice*) seems to refer to the *form* of the verb, i.e. the way a word *sounds*. It is, indeed, a translation of the Greek term φωνή which means 'sound' as well as 'voice' (in the physical sense). However, it could also be used by the grammarians in a more technical fashion as in the sequence of sounds that form a word. The meaning of the Greek grammatical term διάθεσις is very hard to pin down. For valuable discussions of the ancient grammarians on διάθεσις and μεσότης (middle), I refer to Collinge (1963), García Gual (1970), Rijksbaron (1987), Anderson (1989, 1993), and Lallot (1989, 1997).

[20] The term originated from the narrower, emotional meaning 'affect', which appears especially in the German tradition. For instance, Schwyzer (II: 228) emphasizes that "das affektische Moment" is crucial to the semantics of middle verbs. The verba affectuum (verbs of affection, e.g. ἥδομαι, φέβομαι) are a special and productive group, of which the active transitive counterparts are either rare or non-existent. Therefore, Schwyzer concludes, the 'affective moment' must also be the distinctive feature in opposition to the active in the other verb classes. This conception of 'affectedness' is obviously too narrow. The difficulties are especially apparent in the case of an inanimate subject (e.g. ὁ σίδηρος τήκεται 'the iron is melting').

notion πάθος pertains to *passivity*[21], as opposed to ἐνέργεια, which pertains to *activity*. The ancient grammarians, however, are clear in that they consider indirect reflexive *middle* verbs such as ἐποιησάμην, ἐγραψάμην as having an ἐνέργεια-meaning.[22]

Lyons' definition was taken up by Barber (1975). She presents the semantic implications of the active and the middle in Greek in the following diagram (Barber 1975: 21):

Fig. 3 ACTIVE vs. MIDDLE (Barber 1975)

ACTIVE (no incoming arrows)		MIDDLE (incoming arrows)	
Plain Active	SS →	Full Middle	SS ⇄
		Reflexive	SS ↩
		Reciprocal	SS ⇄
		Passive	SS ←

The letters 'SS' stand for Sentence Subject. The arrows indicate the direction of the action resulting in affectedness. The term 'Full Middle' corresponds roughly to the indirect reflexive middle. Barber adds that "(...) the active represents the one case in which the subject is not specified as being affected by the action; whereas the middle subsumes all cases in which the subject *is* affected by the action - (...)" (Barber 1975: 21). In sum, Barber sees the middle voice as a marker of *subject-affectedness*, taken in the broad sense. This comprises, on the one hand, affectedness in which the subject is very much like a patient (as in the passive, reflexive and reciprocal middle), and, on the other hand, affectedness in which the subject is similar to an indirect object (as in the indirect middle).

---

[21] The terms πάθος and παθητική are applied by the ancient grammarians not only to the passive in the strict sense (e.g. τύπτομαι 'I am being beaten'), but also to direct reflexives (e.g. ἐλουσάμην 'I washed myself' (Apollonius, Macrobius and Choeroboscus; see Rijksbaron 1987: 434n8, Lallot 1997, I: 216), and even to *active* verbs like νοσῶ 'be ill' and ὀφθαλμιῶ 'suffer from ophthalmia' (Macrobius; see Rijksbaron 1987: 438). Furthermore, Apollonius Dyscolus (150 Uhlig) has a subclass of παθητική called αὐτοπάθεια 'auto-passivity', which includes active verbs like θνήσκω, ὀφθαλμιῶ, πάσχω (see Lallot 1997, I: 255). In sum, it appears that πάθος and παθητική are *semantic* terms, designating events in which the subject undergoes the action. This means that the term πάθος only partly relates to the meaning of the middle voice. Incidentally, according to Lallot (1997, II: 246n359), the element αὐτο- in αὐτοπάθεια expresses that the element of passivity is designated by the verbal lexeme *itself* (i.e. not by means of passive morphology).

[22] This seems to be implicated by Apollonius III, 30 (see Lallot 1997, I: 216; II: 172). In the ancient grammarians, the term μέση/ μεσότης (middle) seems to refer to a formal category, consisting of (i) 'passive' forms with active meaning (e.g. ἐποιησάμην), and (ii) active forms with passive meaning (e.g. γέγονα). The term could also apply to those forms that can both have an active meaning and a passive meaning, e.g., βιάζομαι 'I assault s.o.' and 'I am being assaulted', διέφθορα 'I have destroyed' and 'I am destroyed' (cf. Dionysius 13.8; Lallot 1989: 55, 166-7).

In this study, I will subscribe to the notion of *subject-affectedness* as the abstract meaning of the middle voice. The advantages are that it is not too vague, and that it subsumes the different middle uses in an adequate fashion.

I would now like to discuss the question as to how the notion of prototypical transitivity is relevant to the semantics of the middle voice. As I have expounded in section 1.1.1, the prototypical transitive clause can be defined in the following manner: an agent-subject volitionally initiates physical activity resulting in a transfer of energy to a patient-object that absorbs the energy and thereby undergoes an internal change of state. As a rule, the verb in a prototypical transitive clause has the active voice. Now the middle voice can be defined as a marked coding of a departure from the prototypical transitive. Contrary to the prototypical transitive, the subject, in some way or other, undergoes an effect of the event. This effect can be of a physical or a mental nature, and it can be direct or indirect (in that it involves an external object). In chapter 3, it will be argued that the notion of prototypical transitivity is also relevant to the semantics of the active-middle-passive trichotomy in the aorist stem.

1.1.4 Markedness and the Meaning of the Active Voice

In the previous section, I have been concerned with the abstract meaning of the middle voice. The meaning of the middle voice was characterized as a marked coding of a departure from the prototypical transitive event. Now the question arises as to the meaning of the active voice. Can we simply conclude that the active voice designates the *absence* of subject-affectedness in opposition to the middle voice? I will argue here that this conclusion cannot be drawn. Instead, the active voice must be taken as the unmarked member of a privative opposition. In other words, the active voice is neutral as to the semantic feature of subject-affectedness.[23]

A useful approach to markedness-phenomena can be found in Croft (1990). Croft argues that the many criteria that have been proposed (notably by Greenberg (1966)), can be reduced to three general ones (cf. 1990: 64ff.):

1 **Structural**: number of morphemes used to express marked and unmarked values.
2 **Behavioural**:
- (a) **Inflectional**: number of cross-cutting distinctions the marked and unmarked values contain.
- (b) **Distributional**: number of syntactic environments in which the marked and unmarked values occur;
- (c) **Cross-linguistic**: number of language types in which the marked and unmarked values occur.

3 **Frequency**:
- (a) **Textual**: number of occurrences of the marked and unmarked values in text;

[23] The markedness of the middle voice in Greek has been advocated earlier by García Gual (1970: 11-12, 29-32), and Ruipérez (1986). Gonda (1979: 39) reaches the same conclusion with respect to the middle voice in Rigvedic Sanskrit, which is highly similar to that of Ancient Greek: "(...) the active is not essentially the exact opposite of, or contrary to, the medium, but it is characterized by not expressing that which is conveyed by the medium, viz. some special reference to the subject. That means that it can be used when the author [of a Sanskrit text, *RJA*] does not deem it necessary to express the medial modification explicitly".

(b) **Cross-linguistic**: number of languages in which the marked and unmarked values are found.

I will argue now that these markedness criteria support the claim that the active voice is the unmarked category, whilst the middle voice is the marked category. Note that two criteria relate to cross-linguistic comparison (2c, 3b). These criteria I will leave out of consideration.[24]

The criterion of *structural markedness* (1) involves counting the morphemes that signal the category in question. Croft (1990: 73):

(24) ***Structure*: The marked value of a grammatical category will be expressed by at least as many morphemes as the unmarked value of that category.**

Compare the regular thematic endings (plus thematic vowel, since they are not always clearly separable) of the present stem indicative in Attic Greek:

Table 2: *Active and middle: present and imperfect indicative endings*

	Present		Imperfect	
	Active	Middle	Active	Middle
1 sg	-ω	-ο-μαι	-ο-ν	-ο-μην
2	-εις	-ηι	-ε-ς	-ου
3	-ει	-ε-ται	-ε-Ø	-ε-το
2 du	-ε-τον	-ε-σθον	-ε-τον	-ε-σθον
3	-ε-τον	-ε-σθον	-ε-την	-ε-σθην
1 pl	-ο-μεν	-ο-μεθα	-ο-μεν	-ο-μεθα
2	-ε-τε	-ε-σθε	-ε-τε	-ε-σθε
3	-ουσι	-ο-νται	-ο-ν	-ο-ντο

Greater morphological complexity of the middle compared to the active is found in the 1st person sg. present (-ω vs. -ο-μαι), 3rd person sg. (-ει vs. -ε-ται), and 3rd person singular of the imperfect (-ε-Ø vs. -ε-το). In the 1st and 3rd sg. present, the active form contains a unanalyzable portmanteau-morpheme (-ω and -ει) that simultaneously expresses *aspect stem, tense, voice*, and *person*. Conversely, the middle forms contain two morphemes; one expresses *aspect stem* (the thematic vowel, -ο- or -ε-), the other expressing *tense, voice*, and *person*. The fact that a number of middle endings display a greater complexity is evidence that the middle voice is marked as compared to the active.[25]

---

[24] That the 'medio-passive voice' was typologically marked compared to the active voice has already been observed by Greenberg (1966: 45-6).

[25] It is interesting to note that the middle endings in Proto-Indo-European (PIE) are a composite of active (secondary indicative) endings and stative endings. E.g., the 1 sg. middle ending **-m-h_2* consists of the active ending **-m* and the stative ending **-h_2*; the 3 sg. middle ending **-t-o* is composed of the active ending **-t* and the stative ending **-o* (see Beekes 1995: 240-2). Thus, in contrast to the active, the middle endings in PIE are clearly morphologically marked, through the addition of the stative morpheme. This stative

Another way of establishing structural markedness involves comparing the *phonological heaviness* of the respective morphemes. Croft (1990: 76) expresses reservations about the validity of this criterion[26]. Although there is a tendency for morphemes expressing marked values to be physically longer, there is frequently a mismatch between physical length and markedness. The example given by Croft is from Spanish: the 1st person plural of the verb *hablar* 'speak' is *hablamos* 'we speak' vs. 2nd person *habláis* 'you speak'. Here, the fact that the ending of the first-person plural contains more phonemes, conflicts with the typological evidence for person markedness. There is a cross-linguistic tendency for third persons (indicative) to be least marked, followed by first persons, and for second persons to be most marked (see Croft 1990: 93, citing Greenberg). For a clear example of this kind of 'markedness mismatch' in Greek, consider the second-person singular endings of the middle, -ηι (present) and -ου (imperfect). Although there is a tendency for 2nd persons to be marked as compared to the 1st and 3rd persons, the Greek 2nd person present and imperfect endings are shorter than the other persons.

Generally, these mismatches will be the result of a historical-phonological change. One example is the disappearance of a phoneme as a result of a sound-law. In the case of the second person middle in Greek, an intervocalic -σ- has disappeared (*-ε-σαι > -ε-αι > -ηι and *-ε-σο > -ε-ο > -ου). In the Spanish example, an intervocalic *-t-* was deleted (cf. its Latin predecessor *fabulatis*).[27]

It is clear that we must, indeed, exercise some caution when we attempt to determine markedness on the criterion of phonological heaviness. However, when establishing a markedness pattern in the active and middle voices in Greek, we can take a *number* of forms (namely two times eight personal forms, see the table above) into consideration, instead of just one pair (as in the examples above). In this fashion, we may be able to detect a tendency that is more robust.

Thus, when we contrast the active endings with the corresponding middle endings (-ω vs. -ο-μαι, etc.), we can observe that the middle endings contain more phonemes in 14 of the 16 forms of the paradigm. Only in two cases, namely the two second-person singular forms, is the pattern different. The primary endings (-εις vs. -ηι: both syllables

---

morpheme was quite possibly a clitic personal pronoun in origin. In the course of time the original compositionality of the middle endings in PIE had become blurred. For the different theories about the origin of the PIE middle endings and the semantics of the middle voice in PIE, see Narten (1968), Neu (1968a, 1968b, 1976, 1985, 1989), Oettinger (1976, 1993), Jasanoff (1978), Strunk (1980), Kortlandt (1981), Georgiev (1985), Rix (1988), Stempel (1996), Kurzová (1999), and Pooth (2000).

[26] Cf. also Comrie (1976: 111): "(...); in general, morphological criteria are the least telling, since the morphology often reflects systematic correspondences of an earlier period of a language."

[27] Interestingly, the irregular 2nd person middle ending -ηι is later replaced by the transparent ending -ε-σαι. This restoration of the 2nd person ending first occurred in the athematic verbs (e.g. ἵστα-σαι in Homer), analogically after forms like κάθημαι : κάθησαι (where the -σ- resulted from -σσ-). In the Koine, the thematic verbs are affected, e.g., -ποιεῖ-σαι (see Schwyzer, I: 668-9). In this way, the expected markedness pattern is restored. The new middle ending -ε-σαι is morphologically more complex, and phonologically more heavy than the active ending -εις.

have three morae[28]) and the secondary middle endings (-ες vs. -ου: both have two morae) are equivalent as to phonological heaviness. Here again, the exceptional case is due to a sound-change: the form *-ε-σο was originally longer than the correspondent active form -ε-ς.[29] In sum, it is safe to conclude that the middle voice is structurally (in particular, phonologically) marked compared to the active[30].

This brings us to Croft's second criterion: *behavioural markedness* (1990: 77). This criterion can be divided into two types: a morphological type, which Croft calls *inflectional* (2a), and a syntactic type, called *distributional* (2b). The former pertains to the number of forms in an inflectional paradigm. The latter pertains to the number of syntactic contexts in which a grammatical element can occur.

Croft defines the inflectional criterion thus (1990: 79):

(25) ***Behaviour (inflectional)*: if the marked value has a certain number of distinct forms in an inflectional paradigm, then the unmarked value will have at least as many forms in the same paradigm.**[31]

With respect to inflectional behaviour, the active and the middle exhibit an important difference. The active voice has two sets of endings (conjugations), the *thematic* (1 sg. -ω, 2 sg. -εις, 3 sg. -ει, and 3 pl. -ουσι), and the *athematic* (1 sg. -μι, 2 sg. -ς, 3 sg. -σι, 3 pl. -ασι), whereas there is only one set of middle endings ((-ο)-μαι, etc,). The fact that the active voice has more distinct forms is an indication that it is the unmarked category as compared to the middle.

The second type of behavioural criterion discussed by Croft (1990: 81ff.) is the *distributional* type (2b). This relates to the number of environments in which the linguistic element in question occurs:

(26) ***Behaviour (distributional)*: if the marked value occurs in a certain number of distinct grammatical contexts (construction types), then the unmarked value will also occur in at least those contexts that the marked occurs in.**

According to Croft (1990: 89-91), the phenomenon that is often called 'neutralization' or 'neutral value-criterion', is to be considered a subtype of the distributional criterion. *Contextual neutralization* implies that, although the meaning predicts that both forms

---

[28] My measure of morae is as follows: a syllable ending in a short vowel has one mora; a syllable which ends in a long vowel or a consonant has two morae; syllables ending in a long vowel *and* a consonant have three morae.

[29] Like the primary indicative -ηι >> -ε-σαι, in later Greek the ending -ου is analogically replaced by the transparent -ε-σο.

[30] For a similar conclusion regarding Modern Greek, see Haspelmath (1993: 99). Note further that the aorist endings (in Ancient Greek) confirm the markedness pattern. In all persons but the second (-σα-ς vs. -σω (< *-σα-σο)), the middle ending is longer. I did not take the aorist forms into account here because of the complication of the passive aorist.

[31] Greenberg (1966: 29) gives a more specific criterion: *An unmarked form will have at least as many allomorphs or paradigmatic irregularities as the marked form.* In the same vein, Comrie (1988: 19-20) notes that the variety of conjugational classes is larger in the unmarked member of the opposition.

should be able to occur in a given context, in fact only one occurs, this being the unmarked form. In Ancient Greek, it can be observed that in a number of contextual environments, the active voice is used, even though, from a semantic point of view, the occurrence of the middle voice would be considered possible. Consider the following examples:

(i) Active verbs are used suppletively in passive constructions (K-G, I: 98-100; Smyth-Messing 1956: 397-8; Schwyzer-Debrunner, II: 226-7). The clearest example of this phenomenon is the active verb (ἀπο)θνήσκω 'die' which is used as a suppletive passive to (ἀπο-)κτείνω 'kill'. So (ἀπο)θνήσκω ὑπό means 'be killed by'.[32] The active voice cannot possibly express absence of subject-affectedness, since that would be in clear contradiction to the passive semantics of the construction. Other examples of this phenomenon are φεύγω 'flee', used as a passive of διώκω 'prosecute' in juridical contexts. Thus, φεύγω ὑπό means 'be prosecuted by'. In compounds, -πίπτω 'fall' functions as a passive of -βάλλω 'throw', e.g. ἐκπίπτω ὑπό 'to be thrown out by'.

Another example of an active form with a passive meaning is the aorist form ἑάλων 'was caught', which has a middle present form ἁλίσκομαι 'be taken, be caught' (see also section 5.0).

(ii) The active voice may be used in contexts in which it is clear that the subject benefits from the action. Clear examples are those in which active and middle forms are used alternately. If it is inferrable from the context that the action is performed in the interest of the subject, the use of the middle form is not obligatory (Gildersleeve 1900: 66, García Gual 1970: 12). For instance,

(27) [The Peloponnesians] (...) ἀναγαγόμενοι ἅμα ἕῳ ἔπλεον, ἐπὶ τεσσάρων **ταξάμενοι** τὰς ναῦς, παρὰ τὴν ἑαυτῶν γῆν ἔσω ἐπὶ τοῦ κόλπου δεξιῷ κέρᾳ ἡγουμένῳ, ὥσπερ καὶ ὥρμουν· ἐπὶ δ' αὐτῷ εἴκοσιν **ἔταξαν** τὰς ἄριστα πλεούσας, (...) (Th. 2.90.1-2)

(...) so they put out to sea at dawn, and, after lining up their ships four deep, sailed along their own shore towards the inner part of the gulf, in the same order as they had lain at anchor, their right wing leading the way. Upon their right wing they had placed their twenty best sailing ships, (...).[33]

The middle ταξάμενοι expresses that the Peloponnesians line up their ships for their own use. In the case of the active form ἔταξαν, self-profit is left implicit.

(28) a. [The Egyptians] (...) ἀπο ὀλυρέων **ποιεῦνται** σιτία, (...) (Hdt. 2.36.2)

(...) they make their bread from spelt (...).

b. [The Egyptians] ἀρτοφαγέουσι δὲ ἐκ τῶν ὀλυρέων **ποιεῦντες** ἄρτους, (...) (Hdt. 2.77.4)

They eat bread which they make from spelt (...).

Here again, the active form (ποιεῦντες) is used in a context in which it can be inferred from the verb ἀρτοφαγέουσι 'they eat bread' that the bread is made for the benefit of the subject (see also Cock 1981: 16).

---

[32] The regular passive form κτείνομαι does not occur in Attic.

[33] The translations accompanying the examples are taken from the Loeb-editions. In places, they have been adjusted.

Additionally, a comparable type of neutralization is found in the verbs of eating and drinking. This class of verbs are mostly active, e.g. ἐσθίω, βιβρώσκω, τρώγω, πίνω (see Schwyzer-Debrunner, II: 225-6). Since these activities are *inherently* for the benefit of the subject, there is no need to code them with the middle voice. For a similar reason, an expression such as ἀνοίγω τὴν θύραν 'I open the door' is always active. The (indirect reflexive) middle form is never used, because it is only natural that one opens the door for oneself.

(iii) A third type of neutralization occurs in contexts with reflexive pronouns. In the following example, the active voice is used in combination with a dative reflexive pronoun αὑτῷ 'for oneself'. The reflexive pronoun is used - instead of an indirect reflexive middle without pronoun - in cases in which the reflexivity is emphasized (i.e.: 'for him*self*'). In these cases, also the middle voice would have been possible from a semantic point of view, (K-G, I: 110-1):

(29) [Philip weakened Macedonia] (...), ἔτ' ἐπισφαλεστέραν ἢ ὑπῆρχε φύσει **κατεσκεύακεν** αὑτῷ (Dem. 2.15)
[Philip] has rendered [Macedonia] for his own benefit even more insecure than it was by nature.

That also the middle voice would have been possible is shown by the following example, where we find both a middle verb and a reflexive pronoun:

(30) τί τὴν πόλιν, Αἰσχίνη, προσῆκε ποιεῖν ἀρχὴν καὶ τυραννίδα τῶν Ἑλλήνων ὁρῶσαν ἑαυτῷ **κατασκευαζόμενον** Φίλιππον; (Dem. 18.66)
What should the city have done, Aischines, when she perceived that Philip was establishing a tyranny over the Greeks for him*self*?

A similar phenomenon is found in *direct* reflexive contexts: sometimes the active form is used in combination with a reflexive pronoun, in cases in which a middle form would have been possible as well, e.g.:[34]

(31) Ἄδρηστος δὲ (...), οὗτος δὴ ὁ φονεὺς μὲν τοῦ ἑωυτοῦ ἀδελφεοῦ γενόμενος, φονεὺς δὲ τοῦ καθήραντος, (...) **ἐπικατασφάζει** τῷ τύμβῳ **ἑωυτόν** (Hdt. 1.45.3)
But Adrastus (...), he who had become the slayer of his own brother and the slayer of his purifier, (...), slew him*self* over the grave.

The following example shows that the middle voice may also be employed[35]:

(32) καὶ οἱ μέν φασι βασιλέα κελεῦσαί τινα ἐπισφάξαι αὐτὸν Κύρῳ, οἱ δ' ἑαυτὸν **ἐπισφάξασθαι** σπασάμενον τὸν ἀκινάκην (Xen. *An.* 1.8.29)
And one report is that the King ordered someone to slay him upon the body of Cyrus, while others say that he drew his dagger and slew him*self*.

---

[34] Gonda (1979: 51) mentions that the same rule (i.e. when a reflexive pronoun is present the middle form is optional) applies for Rigvedic Sanskrit. This rule was already articulated by Pānini (1.3.77).

[35] The direct reflexive middle -σφάττομαι without reflexive pronoun expresses 'to kill oneself' without emphasis or contrast (see also section 2.1.7).

In both cases mentioned above, the reflexive pronoun is used emphatically in a contrastive context (see also section 2.1.7).

(iv) Another indication that the active voice is the unmarked value relates to the passive aorist form (see also Ruijgh 1991: 691). If we judge the passive aorist form solely by its endings (1 sg. [-θη]-ν, 2 sg. -ς, 3 sg. -Ø, etc.) it should be regarded as belonging to the active voice. However, in the system of voice oppositions, in the aorist stem the suffixes -σα- and -(θ)η- also play an essential role. Therefore, it is justified to take the passive aorist as a distinct voice. The active endings are used unproblematically in the passive aorist, since the 'passive' meaning (i.e. subject-affectedness) is expressed by the portmanteau morpheme -θη- which codes both aspect and voice. In my view, the case of the passive aorist forms clearly demonstrates the semantic neutrality of the active voice endings. Apparently, active endings can even be used in contexts of high subject-affectedness such as the passive aorist.[36]

As opposed to the passive aorist formations, the passive future forms do show middle endings (λυθήσομαι 'I will be made loose'; not: *λυθήσω). This redundant marking of subject-affectedness can be explained analogically with pairs like ἔβην - βήσομαι, ἔστην - στήσομαι, and perhaps under influence of the middle futures with passive meaning (e.g. αἱρήσομαι 'I will be taken'). The redundancy of the middle endings in the passive future in Attic and other dialects, is shown by the fact that the Doric dialects have active endings in the passive future (-θη-σέω).

To conclude the discussion of distributional behaviour, we have seen that the active voice readily occurs in contexts in which the element of subject-affectedness is clearly present. This, subsequently, entails that the active voice cannot be thought of as a marking of *absence* of subject-affectedness. Rather, the active voice must be considered as *neutral* to the element of subject-affectedness.

Croft's next criterion regards text-frequency (3a) (1990: 85):

(33) ***Frequency (textual)*: if a marked value occurs a certain number of times in frequency in a given text sample, then the unmarked value will occur at least as many times in a comparable text sample.**

The table in Rijksbaron (1994: 139) shows that the middle voice is, indeed, less frequent. In Herodotus 39 % of the counted verb forms were middles; in Plato 33 %. The ratio active versus middle forms in a given text is approximately 2 : 1. The slight difference between Herodotus and Plato is largely due to a difference in the use of the passive. As is well known, the use of passives is dependent on genre-factors.[37]

---

[36] In ch. 3, I will argue that the passive aorist form conveys a high degree of subject-affectedness. A contrary view is advocated by Bakker (1994). He sees the active endings of the passive aorist as designating the *absence* of affectedness which is due to the punctual aspect of the aorist stem: "Now when the passive event is construed, not in its inherent duration, but as an objective punctual event (e.g. punishment as the transitional point between the state of being guilty and the state of having been punished [as in the case of κολασθῆναι, *RJA*]), the affectedness disappears, and this is signaled by *-thē-* + active (=non-middle) morphology)" (Bakker 1994: 40).

[37] An interesting example of this phenomenon is the fact that passives are fairly rare in the oral narrative of Homer (e.g. the passive of δίδωμι appears only twice, against hundreds of active forms), whereas passive verbs are well-attested in the administrative Mycenean texts (e.g. passives of δίδωμι: *di-do-to*, *de-do-me-na*).

To summarize, all markedness-criteria point in the same direction: the active voice is the unmarked member in the opposition with the marked middle voice. Since the active voice can occur in environments in which the subject is affected (contextual neutralization), it can be concluded that the active is unspecified as to the semantic feature *subject-affectedness*.[38] Conversely, the middle voice is semantically marked with respect to affectedness of the subject. As a consequence, event types that do not involve subject-affectedness cannot be expressed by a middle verb. The two major event types that do not involve subject-affectedness are the *prototypical transitive* and the *stative* event type. In the prototypical transitive event, the subject is an unaffected volitional agent, while the object is the sole participant undergoing the effect of the event. In stative events there is no affectedness, since affectedness can only be the result of a *change* that is taking place or has taken place. Examples of active stative verbs are: εἰμί 'be', ζῶ 'live', καθεύδω 'sleep', μένω 'stay, remain', ὄζω 'smell'. That stative verbs tend to be active is also shown by the denominative verbs meaning 'be NOUN/ADJ', e.g. ἀσθενέω 'be weak' (ἀσθενής 'weak'), βασιλεύω 'be king' (βασιλεύς 'king'), δουλεύω 'be a slave' (δοῦλος 'slave), ἐρυθραίνω 'be red'[39] (ἐρυθρός 'red'), εὐτυχέω 'be prosperous' (εὐτυχής 'prosperous'), ἡσυχάζω 'be quiet' (ἥσυχος 'quiet').

1.2 Polysemy and the Middle Voice

In section (1.0), Rijksbaron's distinction of five different middle usage types was presented. In chapter 2, I will argue that it is useful to distinguish even more middle uses. On the other hand, as we have seen above, it has been the *communis opinio* that it is possible to assign an abstract meaning to the middle voice. How do these two facts relate? The question is whether we should pursue a *polysemous* approach, that is, attempt to distinguish the different specific meaning variants, and to analyze the way they are interrelated. Or, should a *monosemous* approach be pursued by trying to define a core-meaning that is shared by all middle uses? In the latter, the different middle uses are to be considered no more than variant *interpretations* that are completely determined by the context.[40] An objection to the monosemous approach in general is that definitions of the abstract meaning of lexical items or grammatical categories tend to be unspecific to such an extent that also the meaning of oppositional forms (in our case active forms) are covered by them.

However, although one can raise some objections to the monosemous approach, it is indeed possible to define an abstract meaning that is inherent to all variant middle uses. As we have seen in section 1.1.3, this abstract meaning can be characterized as *subject-affectedness*. If we dispense with this observation, we would miss an important generaliza-

---

[38] The asymmetry of the contrast between active and middle voice can be compared to the contrast between unmarked *lion* and marked *lion-ess*. The former term can also be employed in contexts in which the contrast is neutralized, as in *We saw some bored lions in the zoo*, where the group of lions may well contain some female specimens (see also Cruse 2000: 173).

[39] Note, however, that active ἐρυθραίνω can also mean 'make red'. Compare, by contrast, middle ἐρυθραίνομαι '*become* red'.

[40] With respect to the middle voice in Ancient Greek, this is the position taken by Andersen (1993).

tion. In the following section, it will be argued that the monosemous and the polysemous approach are not necessarily mutually exclusive. An attractive synthesis is offered by Langacker's theory of *complex network categories*.

1.2.1 Langacker's Complex Network Category

The *complex category model*, as it is developed by Langacker (1987), builds on Rosch's psycholinguistic work on semantic categories (e.g. Rosch 1973, 1975, 1978). Rosch's psycholinguistic experiments have shown that semantic categories tend to have a highly complex internal structure. Boundaries between categories are of a flexible and graded nature, and some members are 'better' members than others. Membership of a category is determined according to the degree of resemblance to a central member, or *prototype*. This way of determining category membership is thus radically different from the 'Aristotelian' way which requires a member of a category to possess some *essential attribute*.

Now what is a category prototype? The prototype of a category is generally defined as the *best exemplar* or the *typical instance*. Prototypical members have the largest number of attributes in common with other members of the category and the smallest number of attributes which also pertain to members of neighbouring categories. For example, as was shown by Rosch (1975), the most typical member of the category BIRD turned out to be ROBIN (at least, for North American students, who were her test-subjects). On the category boundary were birds like OSTRICHES (which do not fly) and PENGUINS (which do not fly and do not have clearly distinguishable feathers). Even BATS were sometimes included in the BIRD category, which shows the gradedness and flexibility of category boundaries.[41]

The view that categories are non-discrete can also be found in Wittgenstein's *Philosophische Untersuchungen* (1999: 277-8). In a passage on the category SPIEL, he reaches the conclusion that there is no single property that is common to all games. Rather, the category is structured as a network of overlapping similarities ("ein kompliziertes Netz von Ähnlichkeiten"), which Wittgenstein called "Familienähnlichkeiten" (family resemblances). The category SPIEL is like a family. It is impossible to point to a feature that is shared by all family members. Instead, what one can say is that member A has a certain feature *x* in common with member B, member B shares with member C a feature *y*, and so forth. A notable difference between Wittgenstein's category and the category as it is viewed in prototype theory is the idea that, in the latter theory, not all members are equal: some members are 'more member' than others.

Rosch's idea of categorization as hinging on the perceived resemblance of category members to a prototypical member has inspired a number of cognitive linguists to develop models of categorization. Important examples are the *radial network* model (Lakoff 1987), the *complex network* model (Langacker 1987), and the *family resemblance network* (Taylor 1989). These models are concerned with all linguistic categories, including both *lexical items* (e.g. German *Spiel*) and *morphosyntactic structures* (e.g. the transitive construction).

---

[41] A nice example of how peripheral members of one category can switch to another category is Dutch *wal-vis* 'whale'. The name shows that this animal once was considered a member of the category FISH (Dutch *vis*), rather than a mammal. Clearly, of both categories, FISH and MAMMAL, the whale is a peripheral member.

Linguistic categories typically display a polysemous structure of interrelated and conventional values, centered around a prototype.[42]

Now, Langacker's model differs from the two other cognitive models in one important respect. Besides the level of the concrete, polysemously interrelated meanings of a linguistic unit, he posits a higher-level *abstract schema*. The notions *prototype* and *(abstract) schema* are contrasted by Langacker in the following way:

(34) "A prototype is a typical instance of a category, and other elements are assimilated to the category on the basis of their perceived resemblance to the prototype; there are degrees of membership based on the degree of similarity. A schema, by contrast, is an abstract characterization that is fully compatible with all the members of the category it defines (so membership is not a matter of degree); it is an integrated structure that embodies the commonalities of its members" (Langacker 1987: 371).

The notion of abstract schema is roughly comparable with the traditional notion of *abstract meaning* or *core meaning*. Langacker's *complex network model* can therefore be considered a synthesis of a purely *polysemous* approach (involving one form with different, yet related meanings) with a *monosemous* approach (involving one form with one meaning).

The essential building blocks of Langacker's complex category model are *prototype*, *extension* and *abstract schema*. In order to grasp the idea of a complex category network, it may be useful to have a look at an example of a lexical category given by Langacker himself (Langacker 1987: 373-385).[43] Consider a child in the process of learning the various senses (conventional usages) of the word *tree*. In his early experience, the word is first applied to familiar specimens like oaks, elms, and maples. These concrete applications of the word *tree* may be called *usage events* (Langacker 1987: 66). A usage event can be characterized as an actual instance of language use, involving a full, context-dependent (and therefore richly detailed) understanding, that is paired with an actual vocalization. On the basis of a series of usage events, the child will extract a conception that embodies the commonalities of these trees, while properties that vary from one instance to the next will be ignored. This abstraction of the concept $TREE_1$[44], though it excludes subtler and more contingent properties, will nevertheless be fairly concrete and specific as the concept probably involves intrinsic, characteristic, and cognitively salient properties, such as shape, size, color, brachiation, and leaves. Since this concept of $TREE_1$

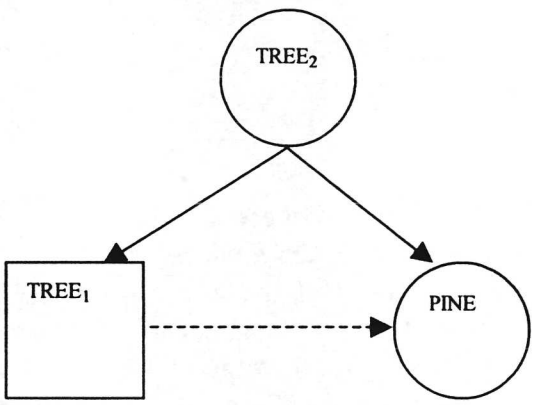
[42] Important works on linguistic prototypes are Givón (1984, 2001a) who discusses the cognitive and biological basis of linguistic prototypes; Lakoff (1987), who discusses the development of the notion prototype in psychology and linguistics, and presents some convincing prototype-analyses; Langacker (1987, 1991a, 1991b); Taylor (1989), on prototype categories; Croft (1991), on syntactic categories; Geeraerts (1993), on lexical semantics; and Goldberg (1995), on constructions. Prototype-based studies on the middle voice are Croft, Shyldkrot & Kemmer (1987), Kemmer (1993), Maldonado (1999), and Manney (2000).

[43] I refer to the passage cited for a considerably more elaborate treatment of the example. For a comparable discussion of the expansion of a lexical category, see Langacker (2000).

[44] Words in capitals here represent *conceptual content* (roughly the *signifié*), whereas small italics represent the phonological content (*signifiant*).

is an abstraction from varied concrete usages, it constitutes a low-level *abstract schema*. With repeated usage, this conception will become more deeply entrenched.[45] This initial conception can be recognized as the category prototype. Subsequently, suppose the child is confronted with a *pine*. He will easily learn to call it a tree, either because somebody calls the pine a tree, or because it is the most applicable term at his disposal. Thus, the original prototypical use of the word *tree* is extended to a new use, namely PINE. This semantic extension is based on the judgement of similarity of the new concept PINE with the highly salient and highly entrenched concept $TREE_1$ which was already acquired by the child. The observation of similarities between the older concept $TREE_1$ (roughly a broad-leaved tree) and the new concept PINE takes the form of an abstract concept $TREE_2$ that embodies the commonality of $TREE_1$ and PINE. This concept $TREE_2$ abstracts away from the conflicting properties of the two lower level concepts, e.g., it will have to be neutral as to the difference between leaves and needles. This higher level concept $TREE_2$ is the *abstract schema* of which the two more concrete uses of the word *tree* are *elaborations*. The process described here results in a simple schematic network depicted in figure 4:

Fig. 4 Extension and Schematization: *tree*



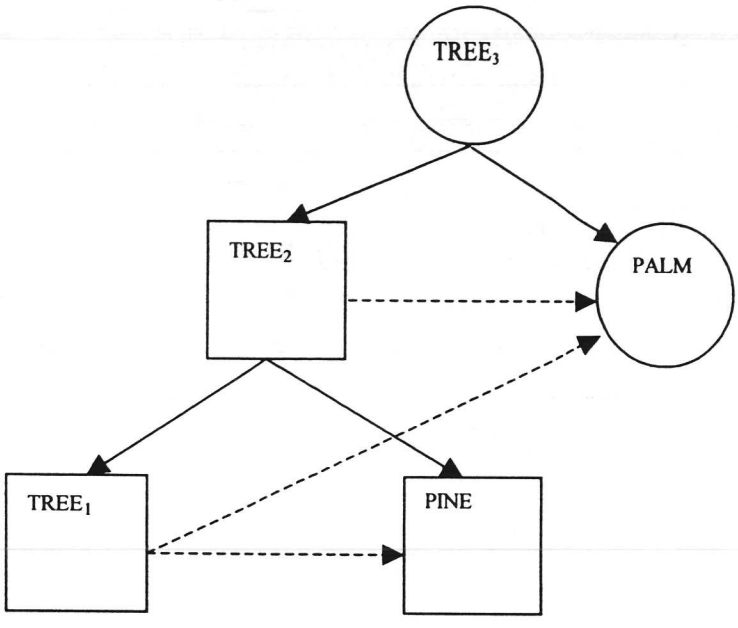
Note that the concept $TREE_1$ is depicted by a square. This is a notational convention used by Langacker to indicate that the concept at issue is *entrenched*. The circles indicate that these concepts are not (yet) entrenched. The solid arrows indicate that the low-level uses are elaborations/specifications of the abstract schema. The dashed arrow indicates that PINE is an *extension* of the prototype $TREE_1$.

The process described above can now be repeated. For instance, the next stage could involve the acquisition of the concept of PALM as a variant use of the word *tree*, which, in turn, results in an even more abstract schematic concept $TREE_3$. This abstract

[45] Every single use of a linguistic structure (e.g. a lexical item or a morphosyntactic structure) has a positive effect on its degree of entrenchment. Conversely, disuse has a negative impact. With repeated use, a novel structure will become progressively entrenched, and it will become a conventional *unit*. A unit is a cognitive structure mastered by the speaker so that it can be employed in a largely automatic fashion, without requiring attention to its individual parts or their arrangement. Since entrenchment is dependent on frequency of use, it is a matter of degree (cf. Langacker 1987: 59-60).

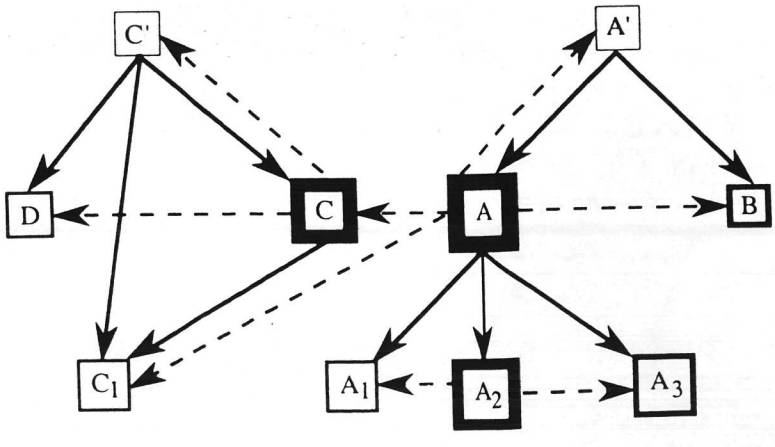
schema should be neutral with respect to, for example, the degree of brachiation, and the size and nature of the foliage. This expanded network category is depicted in figure 5:

Fig. 5 Expansion of the network *tree*



Earlier in this section, the major ingredients of the complex category model were presented, namely *prototype* and *extension*, *abstract schema* and *elaboration*. The complex category can be thought of as a network. The nodes of the network consist of particular linguistic structures. These might be allophones of a phoneme, variant senses of a lexical item (as in the *tree*-example), or variant forms of an elaborate grammatical construction. In addition, the nodes may vary as to the degree of specificity of their content. The nodes are linked to each other in pairwise fashion, thereby indicating their relatedness. A full-blown network category may look like the following diagram (from Langacker 2000: 14):

Fig. 6: *Complex Category Network*



The prototype of the category represented here is node A: almost all the other nodes are either extensions of A (indicated by the dashed arrows), or elaborations (the solid arrows). Node A is also the most entrenched node, which is indicated by the thickness of the box. Node C could be considered a secondary prototype. Furthermore, three levels of schematicity are represented in the diagram. On the lowest level are the structures that are specified in most detail: C_1, A_1, A_2, A_3. On the mid-level the structures are located that are more schematic than the lower-level structures: A and C are abstract schemas of which A_1, A_2, A_3 and C_1 are elaborations (note the solid arrows). The highest level consists of structures A' and C'. These are 'superschemas' of the category. In turn, the mid-level subschemas are elaborations of these superschemas (hence the solid arrows). Another dimension that is represented in the diagram is the aspect of "distance". The nodes are depicted at greater or smaller distances from one another, depending on the difference in degree of elaboration or, in the case of extension, on the number of salient features two nodes have in common.

It is important to keep in mind that the image of a network is a metaphor. It can capture (and visualize) a number of essential properties of the structure of a complex category, as was explained above, but, just like any metaphor, it has its limitations. For example, the discreteness of the boxes in the network may create the false impression that the boundaries between the nodes in the network are clear-cut. In fact, the transition between the nodes (e.g. semantic variants of a lexical item) must be viewed as gradient. In a complex category, there are typically many cases which, having features in common with two nodes (or even more than two), fall between the two of them. Nevertheless, the image of a network can be considered adequate since most instantiations *do* tend to center around a typical case. If this were not so, and in-between cases were more than a relatively marginal phenomenon in our category, we would have to reconsider our analysis, and to see whether we should postulate an additional node in the network.

To improve our understanding of the nature of Langacker's model, it is important to note that it is characterized as a *usage-based model*. Langacker describes this model as follows (1987: 494):

(35) "Substantial import is given to the actual use of the linguistic system and a speaker's knowledge of this use; the grammar is held responsible for a speaker's knowledge of the full range of linguistic conventions, regardless of whether these conventions can be subsumed under more general statements. [It is a] non-reductive approach to linguistic structure that employs fully articulated schematic networks and emphasizes the importance of low-level schemas."

In Langacker (1991a, chapter 10), the model is described as a *maximalist*, *non-reductive*, and *bottom-up* approach, which is consistent with the nature of cognitive linguistics.

Now what do these labels mean? The usage-based model is maximalist and non-reductive in the sense that it recognizes that the linguistic system is a massive, highly redundant inventory of conventional units. In order to become a fluent speaker, a human being has to do a vast amount of actual learning. As Langacker (2000: 2) puts it: "If one aims for psychological reality, it cannot be maintained on purely methodological grounds that the most parsimonious grammar is the best one." Related to this issue is the non-reductive aspect of the model. This concerns the idea that a grammar includes both general rules or schemas as well as expressions that instantiate these rules. For instance, a speaker typically will have learned a general rule for the formation of plural nouns (N + -

s) as well as certain specific instantiating expressions such as *dogs*, *houses*, *cars*. In some approaches to grammar these instantiating expressions are excluded from the grammar on grounds of economy, since these expressions are regularly derivable by rule: to list them individually would be to miss a generalization. This reasoning can be called the *rule/list fallacy* which is based on the assumption that rules and lists are mutually exclusive (Langacker 1987: 29, 2000: 3). Therefore, abstract schemas *and* specific expressions are part of the grammar, provided that they have become conventional units, through entrenchment by frequent, repeated occurrence. As was shown before, abstract schemas are representations of commonalities observed across a number of specific expressions. For instance, the abstract schematic pattern of noun-pluralization is learned through the observation of specific forms like *dogs*, *houses*, etc. These specific expressions must have become entrenched units *before* the abstract pattern is extracted. It is implausible that these specific plural forms will have suddenly lost their status as entrenched, conventional units the moment the general rule is acquired, and that from that moment onward the specific expressions must be computed from scratch on the basis of the rule (cf. Langacker 1991a: 263). It is more plausible that abstract schemas/general patterns *and* specific instantiating expressions coexist as stored units.[46]

I will now move on to the last characteristic of the usage-based approach: the bottom-up orientation. This basically amounts to a difference in emphasis that is given to the different levels of a complex category. Low-level schemas, which are extracted when a language user comes into contact with a fully contextualized instance of a category, are considered of more importance in language processing than higher-level schemas that are extracted from the low-level schemas. "If high-level schemas are extracted, they may be only of secondary significance, serving more of an organizing function than an active computational one" (Langacker 1991b: 265). For instance, when somebody hears the word *dogs*, he will immediately access from memory the highly entrenched low-level unit *dogs* as a pre-packaged whole, instead of computing its structure on the basis of the higher-level schema of noun-pluralization (N + *-s*). Moreover, many complex categories, both lexical and grammatical, lack a single abstract superschema that subsumes all the lower-level values shown by a particular form. In that case, we are dealing with a classic instance of *family resemblances* (as for instance Wittgenstein's SPIEL).

The view that lower-level schemas are more important to language structure than high-level schemas is in accordance with the observation that it is generally not possible to predict the actually occurring instantiations of the high-level schema. This entails that the language user has to know which of the potential low-level instantiations of the high-level schema actually do exist, and which don't. For instance, the passive form *κτείνομαι ὑπό 'be killed by' could, potentially, be sanctioned by the abstract schematic meaning of the middle voice, namely *subject-affectedness*. However, a speaker of Attic Greek had to learn that this instantiation is ungrammatical, and that the construction θνήσκω ὑπό was used instead.

---

[46] The consequence of this view of grammar is that there is no principled division between lexicon and grammar. The grammar of a language is viewed as a *structured inventory of conventional linguistic units*. Linguistic units may vary in degree of compositionality and specificity. The more specific and the less compositional a fixed unit is, the more it can be viewed as a lexical item in the traditional sense. The point is that this is a matter of degree, just as the determining parameters compositionality and specificity.

1.2.2 The Middle Voice as a Complex Category

One of the main objectives of the present study is to show that the grammatical category of middle voice can be insightfully analyzed as a complex network category. In section 1.0, I described the five different middle usage types as identified by Rijksbaron (1994[2]). These are the passive use, the direct reflexive use, the indirect reflexive use, the pseudo-reflexive/pseudo-passive use, and the middle-only use. These usage types can be identified with the concrete, low-level senses of the complex category. These variant uses of the middle voice are semantically related to one another in a polysemous fashion. In the complex category models, these relationships are called *extensions*. This means that, although they share certain semantic features, there are salient features in which they differ.

As was noted above, the abstract meaning of the Greek middle voice can be defined as *affectedness of the subject*. This abstract meaning can be identified with the *abstract schema* in Langacker's model. In accordance with the usage-based model of grammar, it is conceivable that this abstract schema is less entrenched, and only of secondary importance in actual language use. In speaking and hearing, the language user is more likely to activate the more concrete middle usage types, than the rather abstract superschema of *subject-affectedness*. For example, it is plausible that, when a Greek heard the word ἵσταμαι in a context without a direct object or external agent, the low-level "node" of the *pseudo-reflexive*[47], that specified that the subject undergoes a self-initiated change of state, was activated first and foremost. The abstract schema, with the single implication that the subject is affected, may have been activated less strongly, or not at all.[48] Obviously, this assumption cannot be tested in a dead language like Ancient Greek, and it is, therefore, not crucial to the purposes of my argument.

In chapter 2, I will give an overview of the different middle usage types which I distinguish, and I will describe the structure of the semantic network of the middle voice. In several respects, my classification will be finer-grained than Rijksbaron's presented above. This refinement is primarily supported by typological evidence (esp. Kemmer, Croft, Shyldkrot 1987; Kemmer 1993), and by morphological evidence from the aorist stem. After this classification of middle uses, the semantic relationships among them will be discussed.[49] This discussion will result in a "semantic map" of the Greek middle voice. This semantic map represents the precise structure, as I see it, of the semantic network of middle uses and their interconnections.

---

[47] In chapter 2, I will rename this middle usage type "body motion middle".

[48] It can also be assumed that these concrete middle usages are acquired first. The abstract schema is then extracted from the variant usage types on the basis of the semantic commonalities observed (unconsciously) by the language user (see 1.2.1).

[49] These semantic relationships can be identified with the *extensions* of Langacker's model: one polysemous meaning is an extension of a more prototypical one if - although they share one or more semantic commonalities - there is at least one salient semantic property in which they differ. I shall use the more neutral term *relation(ship)*, because it is often unclear which of the connected uses is the more prototypical, and which is the extension. From a diachronical point of view, one can only guess as to the exact way in which the middle voice emerged since there already must have been a full-blown middle voice in Proto-Indo-European. Synchronically, frequency of use can be an argument for prototypicality. In section 2.3.2, I will try to establish the category prototype of the middle voice in Greek.

The semantic map approach and its implications for distributional patterns and diachrony I owe to Anderson (1982) [on the perfect]; Haspelmath (1987) [on reflexives]; Croft, Shyldkrot & Kemmer (1987) [on the middle voice]; and Kemmer (1993) [on the middle voice][50]. These are all typological studies which attempt to establish a universally valid *semantic map*[51] for each of the grammatical categories concerned. The semantic map of the middle voice in Ancient Greek, which will be based on a semantic analysis of the different middle usage types, will serve as a basis for two claims that will form the point of departure in chapter 3. Both claims relate to morphology, the first from a synchronical point of view, the second from a diachronical point of view:

(I) **A form will always cover a connected region of variant middle uses in the semantic network.**[52]

(II) **A form will only spread from one variant use to another when these uses are directly semantically related.**[53]

These claims are primarily relevant to the analysis of the complex morphology of the aorist stem, which will be the objective of chapter 3. The first claim is based on the general functionalist assumption that different meanings that are expressed by the same form are related, until proven otherwise. With respect to the morphology of the aorist stem, it predicts that every aorist form - i.e. sigmatic middle, thematic middle, root middle, and passive aorist - will map onto a contiguous portion of the semantic network. The second claim predicts that historical expansion of an aorist form through the network will follow the pathways of the semantic links. In other words, the extending form will not "jump over" from middle use A to middle use C, without affecting the intermediate use B. These two claims can be seen as two sides of the same coin, since the synchronical contiguity of a form in the semantic network results from the diachronical tendency of forms to spread only to semantically adjacent meanings.

---

[50] Kemmer's study will be discussed in more detail in section 1.2.3.

[51] Croft (2001) distinguishes the notion *conceptual space* from that of *semantic map*. The former is defined by Croft as follows: "Conceptual space is a structured representation of functional structures and their relationship to each other" (Croft 2001: 93). Conceptual spaces are hypothesized to be *language-universal*. A semantic map is a map of the distributional pattern of a *language-specific* category on a conceptual space. In his thought-provoking book, Croft shows the importance of the notion of *conceptual space* for the understanding of language-specific grammatical categories. He insightfully analyses the different conceptual spaces that relate to parts of speech, clausal syntactic roles, the voice continuum, and the subordination-coordination continuum. Since the present study is not concerned with a cross-linguistic comparison of voice systems, but is restricted to the Ancient Greek middle voice, employing the term *conceptual space* would not be appropriate. Instead, I will use the term *semantic map* or *semantic network* in reference to the particular polysemic structure of the Ancient Greek middle voice.

[52] This principle is related to Croft's *Semantic Map Connectivity Hypothesis*, which runs thus: "any relevant language-specific and construction-specific category should map onto a CONNECTED REGION in conceptual space" (Croft 2001: 96).

[53] This is essentially Crofts's dynamicized version of the Semantic Map Connectivity Hypothesis: *diachronic changes in the distribution of a construction should follow connected paths in conceptual space* (Croft 2001: 105).

Furthermore, these claims can serve as a powerful empirical tool to test the accuracy of our network structure as it results from a semantic analysis only. For example, if an aorist form does not cover a contiguous region in the network, or if an aorist form does not extend gradually through the semantic links in the network, then the map of our semantic network should be seriously reconsidered, and modified in such a way that violations of the two claims are avoided.

1.2.3 Kemmer, *The Middle Voice* (1993)

In the previous section, a network conception of the middle voice was set forth. The major advocate of this approach to the middle voice has been Suzanne Kemmer. Especially her typological study *The Middle Voice* (1993) - a revision of her doctoral thesis - has been an important contribution to our understanding of middle voice phenomena, and it has also been fundamental to the present study. Therefore, it is more than justified to devote a section of my study to a discussion of her book. This section will focus on those analyses and conclusions in the book that are of importance to the argument of the present study.

Kemmer's *The Middle Voice* is a large-scale typological comparison of middle voice systems in 30 languages. After an introductory chapter, Kemmer starts off (chapter 2) by giving an inventory of the middle situation types that are frequently marked by middle morphology across the world's languages. She enumerates 13 types of middle situations (Kemmer 1993: 16-20)[54]:

(1) **Grooming or body care**: Djola (Niger-Congo) *-pɔs-ɔ* 'wash'; Latin *lavo-**r*** 'wash'; Bahasa Indonesia ***ber-dandan*** 'get dressed'; Old Norse *klæða-**sk*** 'get dressed'; Hungarian *borotvál-**koz***- 'shave'.

(2) **Nontranslational motion**: Kanuri (Nilo-Saharan) *tàn-**t**-în* 'stretch one's body'; Old Norse *snúa-**sk*** 'turn'; Latin *reverto-**r*** 'return'; Classical Greek *trépe-**sthai*** 'turn'; German ***sich** verbeugen* 'bow'.

(3) **Change in body posture**: Djola *lak-ɔ* 'sit down'; Bahasa Indonesia ***ber**-lutut* 'kneel down'; German ***sich** hinlegen* 'lie down'; Guugu Yimidhirr (Australian, Pama-Nyugan) *daga-**adhi*** 'sit down'; Hungarian *emel-**ked***- 'rise, get up'.

(4) **Indirect middle**: Turkish *ed-**in*** 'acquire'; Old Norse *eigna-**sk*** 'acquire, lay claim to'; Classical Greek *kta-**sthai*** 'acquire for oneself'; Latin *apīsco-**r*** 'obtain'; Changana (Niger-Congo, SE Bantu) *ku **ti**-tekela* 'take for oneself'.

(5) **Naturally reciprocal events**: Old Norse *hitta-**sk*** 'meet'; Hungarian *ölel-**kez***- 'embrace'; Latin *amplecto-**r*** 'embrace'; Bahasa Indonesia ***ber**-gumul* 'wrestle'; Guugu Yimidhirr *yirrga-**adhi*** 'converse, agree';

(6) **Translational motion**: Pangwa (Niger-Congo, CE Bantu) *i-nu-xa* 'climb up'; Guugu Yimidhirr *madha-**adhi*** 'climb up'; Old Norse *ganga-**sk*** 'go, leave'; Bahase Indonesia ***ber**-jalan* 'walk, stroll'; Classical Greek *péte-**sthai*** 'fly'; Latin *veho-**r*** 'travel'.

(7) **Emotion middle**: Guugu Yimidhirr *dumba-**adhi*** 'get a shock or fright'; German ***sich** fürchten* 'be/become frightened'; Mohave (Hokan, Yuman) ***mat** iθa:v* 'be angry'; Latin *irāsco-**r*** 'become angry'; Hungarian *bán-**kod***- 'grieve, mourn'.

[54] The middle markers are in bold-face.

(8) **Emotive speech actions**: Latin *quero-r* 'complain', German ***sich beschweren*** 'complain'; Classical Greek *olophūre-sthai* 'lament', Turkish *döv-üün* 'lament'; Sanskrit *krpa-te* 'lament'.

(9) **Other speech actions**: Mohave ***mat*** *kuna:v* 'confess'; Pangwa ***-i-lumba*** 'admit one's guilt'; Latin *fateo-r* 'confess'; Kanuri *àwùlò-t-în* 'be boastful, be a braggart'; Hungarian *dicse-ked-* 'boast'.

(10) **Cognition middle**: Bahasa Indonesia ***ber-pikir*** 'be cogitating'; Old Norse *þykkja-sk* 'think'; Pangwa *-i-sala* 'think over, consider'; Latin *medito-r* 'ponder, meditate'; Mohave ***mat*** *ahay* 'believe'.

(11) **Spontaneous events**: Changana *ku-ti-milela* 'germinate, sprout'; Old Norse *gróa-sk* 'grow'; Bahasa Indonesia ***ber-henti*** 'come to a stop'; Mohave ***mat*** *ičo:* 'become, change into'; Turkish *dinl-en* 'recover'.

(12) **Logophoric middle**: the middle marker appears on verbs of saying, belief, and perception in construction with a complement clause, e.g. Icelandic *þeir sáus-st hlaupa* 'They saw themselves run'.

(13) **Passive, Impersonal, Facilitative middles**: Kanuri *t-úrúk-ìn* 'I am seen' (passive); German *Der Artikel liest* ***sich*** *leicht*; French *Le livre* ***se*** *vend bien* (facilitative).

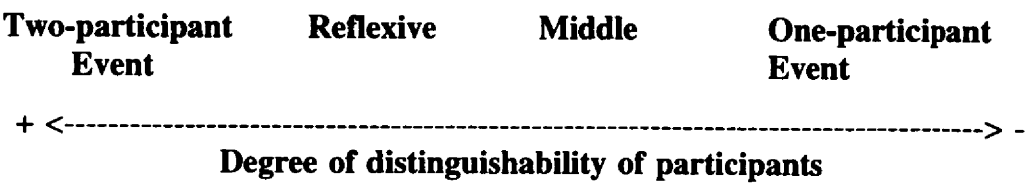
Apart from (12), the *logophoric middle*, and the *facilitative* middle, each of the enumerated middle uses seems to be instantiated in Ancient Greek. In (1) *grooming or body care*, we can recognize the direct reflexive middle use. In (2) *non-translational motion*, (3) *change in body posture*, and (6) *translational motion*, Rijksbaron's (1994[2]) *pseudo-reflexive use*, which almost exclusively consists of verbs of motion, can be recognized. The *indirect middle* (4) can be compared to the *indirect reflexive middle* in Greek. The uses (5) *naturally reciprocal events*, (8) *emotive speech actions*, (9) *other speech actions*, and (10) *cognition middle* are typical media tantum classes in Greek, e.g. μάχομαι 'fight' (5), ὀλοφύρομαι 'lament' (8), μυθέομαι 'speak' (9), λογίζομαι 'calculate; consider' (10). The *emotion middle* (7) and the *spontaneous event middle* (11) correspond to Rijksbaron's (1994) *pseudo-passive use*, e.g. φοβέομαι 'fear' (7) and τήκομαι 'melt (intr.)' (11).

Generalizing from these seemingly very diverse situation types expressed by the middle voice, Kemmer argues that the one crucial semantic property of the middle voice is the *relative distinguishability of participants*. Kemmer (1994: 211):

(36) "Relative elaboration of events can be thought of as the degree to which different schematic aspects of a situation are separated out and viewed as distinct by the speaker. The speaker in effect can choose to "turn up" or "turn down" the resolution with which a particular event is viewed in order to highlight its internal structure to a greater or lesser extent."

This property, according to Kemmer, subsumes the notion of 'subject-affectedness'. Thus, the middle voice can be located on a gradient scale between two extremes, i.e. between two-participant events (high distinguishability) and one-participant events (low distinguishability) (Kemmer 1993: 73):

Fig. 7: *Scale of degree of distinguishability of participants:*



The *two-participant event* is located at the one extreme of the scale. According to Kemmer, in a two-participant event the participants are maximally distinguishable in that the two participants, the Initiator and the Endpoint, are completely separate entities. In a reflexive event the distinguishability is lower. The reflexive marker (the pronoun ἑαυτόν in Ancient Greek) signals the co-referentiality of the participants in events in which the two participants are normally distinct entities. In the reflexive, the separation of the two participants is thus to some extent maintained. In the middle type the dual roles, Initiator and Endpoint, are conflated in a single participant. Thus, the distinguishability of participants is minimal, yet higher than in prototypical one-participant events. In the case of the middle, a certain degree of internal complexity is extant by virtue of the initiating and affected aspect that is evoked.

I already mentioned, in passing, the notions Initiator and Endpoint. According to Kemmer, the notions *Initiator* and *Endpoint* are important for an understanding of middle semantics. To get a grasp of these notions, it is best to consider the transitive clause, which is, according to Kemmer (1994: 51), the basic conceptual model for these notions. We have seen that the prototypical transitive clause involves a physical transmission of energy, initiated by a volitional agent and resulting in a change of state of the patient. However, it is a trivial observation that transitive clauses generally also denote other types of interactions. For instance, in section 1.1.1 I mentioned mental event types that are very often coded as transitive events, even though they involve neither a volitional agent nor a state-changing patient. The coding of these non-prototypical transitive event-types in a morphologically transitive clause is conditioned by the degree to which such events can be construed as an asymmetrical interaction between two participants. In this asymmetrical interaction, the first participant is construed as initiating the event (hence *Initiator*); the second participant is construed as the endpoint of the action chain (*Endpoint*).[55] The notions of Initiator and Endpoint are general semantic roles, subsuming various other, more specific, semantic roles. The Initiator role subsumes those roles that involve a conceptualization of a "starting point" of an event, such as *agent*, *experiencer*, and *mental source*. Endpoint, on the other hand, encompasses the "down-stream" roles, such as *patient*, *recipient* and *beneficiary*. As they subsume the more concrete semantic roles, Initiator and Endpoint are called "macro-roles" by Kemmer.[56] The semantics of the

[55] The notions of Initiator and Endpoint are comparable to Langacker's *head* and *tail* of an action chain (see section 1.1.2).

[56] A related notion, going by the same name of 'macro-role', also features prominently in Role and Reference Grammar (Foley & Van Valin 1984, Van Valin & LaPolla 1997). Klaiman (1988, 1991) has a notion very similar to macro-roles, that of *conceptual status*. Instead of Initiator and Endpoint, she works with the *controller* and *affected entity*. The notion of *controller* is, however, more restricted than that of Initiator, since it roughly only includes animate entities.

middle voice can now be described elegantly in terms of these macro-roles: the middle voice expresses that the subject is conceptualized as both the Initiator and the Endpoint.

With regard to Ancient Greek, however, this characterization of the middle voice is not entirely adequate. As was set out above, the Greek middle voice can be characterized by the notion of *subject-affectedness*. The idea that the subject is affected by the event is roughly equivalent to the idea that the subject is the Endpoint in the event. The notable difference between Greek and many other middle voice systems is that the middle voice in Greek also covers the canonical passive meaning. This means that, in Greek, the subject of a middle verb cannot be conceived of as an Initiator in all cases. This is a crucial difference between the Greek middle voice and many other middle voice systems (esp. those of modern European languages) which originate from reflexive markers.[57] In these languages, the middle category prototype relates to grooming actions (Kemmer 1993: 55), or possibly to the related body action types such as *change in body posture, non-translational* and *translational motion*. The spontaneous event type and the passive[58] are to be considered peripheral usage types. In Ancient Greek, however, the category's gravitational point lies elsewhere. In section 2.3.2, it will be argued that the intransitive middle types[59] - especially the mental process type - constitute the centre of the middle voice in Greek. The grooming type in Greek is to be considered of marginal importance as both type and token frequency are very low.

In chapters 3 and 4, Kemmer gives a thorough description of the different reflexive and middle types found in the languages of the world. As we have seen above, many of these can also be found in Ancient Greek. In chapter 5, the historical emergence of the middle voice from a reflexive origin in a number of language-families (Romance, Germanic and Nilo-Saharan) is discussed. Kemmer describes the spread of the reflexive marker *sē* in Latin through the middle semantic domain in French and Surselvan (a Rhaeto-Romance language), accompanied by the disappearence of the older middle marker, the inflection in *-or, -ris, -tur*, etc. Kemmer describes the expansion of the reflexive marker *se* as a process of grammaticalization: the meaning shifts from an originally emphatic direct reflexive sense in Latin to a more abstract middle meaning in the Romance languages. Eventually, the marker could also be used as a passive. This semantic generalization process is accompanied by typical symptoms of grammaticalization: cliticization (pan-Romance), erosion (loss of phonological substance), affixation (in Surselvan), and the spread of *se* from the 3rd person to all other persons (in Surselvan).

The description of this diachronic process, involving a reflexive marker developing into a middle marker, as it occurred in Romance and Germanic languages, is not directly relevant to the Greek middle voice and its origins. If the Indo-European middle voice did

---

[57] The same observation holds for Latin, and perhaps more ancient Indo-European languages. In Latin, most of the verbs in *-or, -ris, -tur*, etc. have a passive meaning.

[58] Note that in many modern European languages the prototypical passive is expressed by an auxiliary plus participle, e.g. French *La porte est ouverte par Roger.*

[59] Although, strictly speaking, the direct reflexive middle is also intransitive, I will use *intransitive middle* as a cover-term for the body motion middle (roughly Rijksbaron's *pseudo-reflexive*), the mental process middle and the spontaneous process middle (Rijksbaron's *pseudo-passive*). The intransitive middle is called *decausative* by some authors (e.g. Geniušiene 1987, Gerritsen 1990), since they systematically correspond to active causative verbs.

develop from a reflexive marker, it must have been in an earlier stage of Proto-Indo-European (PIE), since we can already reconstruct a reflexive marker for PIE, **s(u)e*, alongside a full-blown inflectional middle voice category (1s **-mh$_2$*, 2s **-sth$_2$o*, 3s **-to*, etc.). In other words, in PIE we find a situation in which the (emphatic) reflexive is expressed by a marker (**s(u)e*) that is not genetically related to the middle inflection (1s **-mh$_2$*, etc.).[60]

However, for students of the Greek middle voice the primary importance of Kemmer's description of the spread of the reflexive marker from one use to another lies elsewhere. The first point of importance is of a methodological nature. Kemmer's diachronical analyses are based on the assumption that two meanings A en B are related, if one can observe that a form spreads from meaning A to meaning B (1993: 5). This makes diachronic evidence as important to a semantic analysis as synchronic evidence. In section 1.2.2, I have expressed my adherence to this principle, and in chapter 3, I will argue that this principle can improve our understanding of the morphological distribution of the different aorist formations and its semantic implications. The second point of importance concerns the historical development of the reflexive marker. Kemmer describes the frequently occurring development from a reflexive pronoun (used with verbs of grooming, and many verbs of motion) to a more general middle marker that also includes mental (emotional and cognitive) events and spontaneous events. In some cases, this middle marker also develops a passive meaning.[61] If we transpose these finding to Ancient Greek, we can assume a semantic relatedness between different middle uses in the following way: at the one extreme, the reflexive-like uses (the direct reflexive middle and the pseudo-reflexive middle) are located, and at the other, the passive is located. In between these uses lie the emotional, cognitive and spontaneous uses (Rijksbaron's *pseudo-passive*). In chapters 2 and 3, I will produce evidence that this configuration is, in principle, correct, and I will, furthermore, propose a number of refinements.

In chapter 6, Kemmer presents a 'semantic map' which represents the network consisting of the various middle situation types and the semantic relations existing among them (1993: 202). This map of the interrelations between middle types is based on the shared semantic properties that were revealed by the typological and diachronic data that were analysed in the preceding chapter of her study. Since this map is based on a large collection of typological data, it is claimed to have universal validity. In other words, the semantic relations established in the map are claimed to be relevant to all languages. In particular, the set of uses found in any one language is predicted to form a contiguous region on the map: "(...) we would in principle expect the semantic range of a given form to extend only over a set of uses in which each use is directly related to at least one other use in the set" (Kemmer 1993: 222). The diachronic correlate of this prediction is that the spread of a given form is predicted to follow the pathways which are formed by the semantic relations, that is, from one use to another, directly linked, use.

As was said at the outset of this section, Kemmer's book is an inspiring and rich *Fundgrube* for any student of voice phenomena. In the discussion above, it must be noted,

---

[60] The situation we find in Ancient Greek, with both a middle inflection and a historically non-related emphatic reflexive marker (e.g. 3s acc. ἑαυτόν 'himself'), appears to be directly inherited from PIE. Note that ἑ-αυτόν contains the morpheme ἑ- which is a direct descendant from PIE **sue*.

[61] This development has also been described by Haspelmath (1987: 35).

I have restricted myself to those aspects of the work that are of particular interest to my own study of the Greek middle voice.

1.3 Two General Issues

In the preceding sections, I have set out the major problems that will be dealt with in this study, and I have presented the basic assumptions that will underlie my analysis. Now I will address two general issues that are important to an understanding of the middle voice. First, we will consider the status of the media tantum within the middle voice system, and secondly, we will discus the notion of valence reduction as an explanation for middle voice phenomena.

1.3.1 The Status of the Media Tantum

Media tantum are traditionally treated as a distinct class (see for instance Smyth-Messing 1956: 393; Schwyzer-Debrunner, II: 228-9). This approach, however, is not entirely unproblematic. First, it is clear that the class of media tantum is semantically highly diverse. A superficial glance at a selection of media tantum may demonstrate this point: ἀγωνίζομαι 'contend', ἀκροάομαι 'listen', ἄλλομαι 'jump', βούλομαι 'want', γίγνομαι 'be born, become', δέχομαι 'accept', ἐπίσταμαι 'know', ἔρχομαι 'go', εὔχομαι 'pray', λογίζομαι 'calculate', πυνθάνομαι 'learn, hear; inquire'. The diversity within this class is also manifested by the differences in aorist morphology. Some of these verbs may have a sigmatic middle aorist, others a passive aorist, and yet others thematic or root aorists. The rationale behind the distribution of the different aorist formations will be the subject of chapter 3.

A second difficulty in treating the media tantum as a separate class is that, in doing so, two inconsistent criteria are aplied. On the one hand, the various oppositional middle types[62] are distinguished purely on the basis of semantic criteria. For instance, the indirect reflexive middle is distinguished by the semantic feature *subject profits from action*, and the passive middle is defined by the fact that the subject is semantically a patient. On the other hand, the media tantum are distinguished by a completely different criterion, namely the non-existence of an active form. This criterion stands orthogonal to the semantic criterion that is applied to the oppositional middles.

From a methodological point of view, it is preferable to aim at a unified account of both oppositional middles and media tantum. Here and in chapter 2, I shall argue that such a unified account is indeed possible on the basis of purely semantic criteria. For a great number of media tantum it is possible to classify them under the already established middle uses in a straightforward way. Many of these media tantum are semantically very close to oppositional middles. It would, therefore, be highly unnatural to treat them as belonging to a separate class for the sole reason that they do not have an active counterpart. A nice case in point is the medium tantum ἥδομαι 'enjoy' which is semantically very close to the oppositional middles εὐφραίνομαι 'be glad' (active causative εὐφραίνω

[62] The *oppositional middles* are those middle verbs that stand in opposition to an active verb. Two typical examples are: (i) middle παρασκευάζομαι (indirect reflexive, direct reflexive, or passive) vs. active transitive παρασκευάζω, and (ii) middle ἵσταμαι (indirect reflexive, intransitive, or passive) vs. active transitive ἵστημι.

'make glad') and Homeric τέρπομαι 'amuse oneself' (active causative τέρπω 'amuse'). It therefore makes good sense to classify ἥδομαι as a pseudo-passive (in Rijksbaron's terminology).[63] Supporting evidence for the view that media tantum have to be treated on a par with the oppositional middles is based on their morphology. The media tantum generally make the same choice for either the sigmatic middle aorist or the passive aorist form as do the oppositional middles of the same semantic type. For example, ἥδομαι has a passive aorist ἥσθην just like εὐφραίνομαι - ηὐφράνθην. I will go into the morphology of the aorist stem extensively in chapter 3.

In the same way, other media tantum can be classified under the already familiar middle uses. On closer inspection, it becomes clear that most media tantum belong to a restricted number of semantic classes. These are:

(37) *Media tantum: semantic types*

- a. **Media tantum with a beneficiary/recipient-subject**: δέχομαι 'accept, receive', ἐργάζομαι 'work, make', κτάομαι 'acquire'
- b. **Media tantum of body motion**: ἅλλομαι 'jump', ἕπομαι 'follow', ἔρχομαι 'go', πέτομαι 'fly'
- c. **Media tantum of emotion and cognition**: ἄγαμαι 'wonder', ἄχθομαι 'be grieved', βούλομαι 'want', ἔραμαι 'love', ἥδομαι 'enjoy', οἴομαι 'think', πυνθάνομαι 'learn, hear'
- d. **Media tantum of volitional mental activities**: λογίζομαι 'calculate', μηχανάομαι 'contrive'
- e. **Reciprocal media tantum**: ἀγωνίζομαι 'contend', μάχομαι 'fight'
- f. **Media tantum of perception**: ἀκροάομαι 'listen', δέρκομαι 'gaze, look at', θεάομαι 'gaze', ὀσφραίνομαι 'smell', σκέπτομαι 'look'
- g. **Speech act media tantum**: αἰτιάομαι 'accuse', ἀράομαι 'pray', ὀλοφύρομαι 'lament', μαρτύρομαι 'call to witness', μυθέομαι 'speak'

The first three types (a-c) can be readily recognized as being related to oppositional middle types. The media tantum with a beneficiary/recipient-subject (a) can be identified as indirect reflexive middles. This identification is also supported - or at least not contradicted - by the aorist morphology as both the indirect reflexive oppositional middles and the indirect reflexive media tantum consistently have a sigmatic middle aorist, and never a passive aorist, as for example, ἐδεξάμην, εἰργασάμην, ἐκτησάμην. The media tantum designating body motion can be classified as pseudo-reflexive middles on a par with middles like ἀπαλλάττομαι 'depart', ἵσταμαι 'stand up, stand still', πορεύομαι 'go, march'. Lastly, the media tantum designating emotion and cognition can be identified with Rijksbaron's (1994²) pseudo-passive middles (type φοβέω 'frighten' - φοβέομαι 'fear'). Again, aorist morphology suggests that uniting the oppositional middles and the media tantum designating mental processes is justified since both always have a passive aorist

[63] Note that, after Homer, an active causative verb ἥδω was created. I would take this as confirmation that the border between media tantum and oppositional middles is not a very important one, since it is easily crossed.

form, and never a sigmatic middle aorist: e.g., oppositional middles: ἐλυπήθην, ἐφοβήθην, ὠργίσθην; media tantum: ἐβουλήθην, ἤσθην, ᾠήθην.[64]

The last four types (d-g) cannot be compared to the already established oppositional middle uses. Yet I will argue that they should not be set apart completely from the other middle uses; they should, instead, be integrated into the polysemous structure of the middle voice, just like the other (oppositional) middle usage types. In other words, these types of media tantum can be seen as nodes in the semantic network linked to the other, more familiar, nodes. In chapter 2, I will show that the four types at issue are semantically closely connected to the other middle uses, and that the boundaries between them are only a matter of gradiency. In chapter 3, it will be shown that there is also evidence from aorist morphology for the existence of polysemous relations between the media tantum and the oppositional middles.

1.3.2 Valence reduction

Earlier in this chapter I argued that the middle voice as a category can be characterized by a *semantic* feature, namely *subject-affectedness*. There is, however, a substantially different approach in the literature on the middle voice and reflexive verbs. This more syntactically oriented approach sees the middle voice as a marking of *valence reduction*. Valence reduction implies, in brief, that by means of the middle voice morpheme transitive predicates are detransitivized. An early example of this approach is Marguliés (1929-30: 116): "Wenn wir bedenken daß es keine idg. suffixalen Intransitivbildungen gibt, dann scheint es nicht zu gewagt, anzunehmen, daß die ursprüngliche Funktion der Medialendungen die Intransitivierung war." More recent attempts along these lines are Aissen (1982), Dik (1982) [on reflexive verbs], Geniušiene (1987) [on reflexive verbs], Risselada (1987) [on Ancient Greek], Dik (1997b: 9-15).

There are, however, a number of serious objections to this approach. Firstly, in probably all middle voice systems, as in Greek, there is a class of media tantum - middle verbs that do not have an active counterpart. Clearly, these verbs cannot be derived from active predicates by an argument reduction rule. As a possible solution for this problem one could 'condemn' the media tantum to the lexicon. This solution is, however, not very elegant since it means that the attempt at a unified account of the middle voice is abandoned, and that a highly frequent class of middle verbs is basically left unaccounted for. Moreover, this view contravenes the usage-based approach adopted here, since the latter approach claims that storage in the lexicon and the existence of a general grammatical rule are not mutually exclusive (the "non-reductionalist" view). Therefore, though it is clear that media tantum are individually stored in the lexicon, it is possible that they can be subsumed under a more general schema together with the oppositional middles. Furthermore, oppositional middles and media tantum cannot be distinguished by the criterion that the former are grammatical, and the latter lexical, since many frequently occurring oppositional middles are probably stored in memory (i.e. "in the lexicon") as

[64] Rijksbaron, too, remarks that the fact that the agentive media tantum have a sigmatic middle aorist is in accordance with the agentive direct and indirect reflexive uses. Furthermore, the passiva tantum semantically resemble the pseudo-passives of causative verbs. Both are mostly non-agentive and intransitive and have passive aorist forms (see Rijksbaron 1994^2: 156-7).

well.[65] For example, it is improbable that a highly frequent verb form like ἵσταμαι 'stand still, stand up' was computed from scratch by a grammatical derivational rule on the basis of the active form ἵστημι 'make to stand' every single time it was uttered. It is more plausible to assume that the form ἵσταμαι had become entrenched as a fixed unit (because of its high frequency), and was retrieved automatically from memory as a pre-packaged whole, which clearly cuts down on processing efforts.

A second objection to the valence-reduction approach is related to the previous one. It is a fact that most intransitive middles have a higher frequency than their active transitive counterparts. The following table shows token-frequencies of active transitive and middle intransitive verbs[66] in Herodotus (based on Powell's lexicon (1938)):

Table 3: *Frequencies of transitive active vs. intransitive middle in Herodotus*

Transitive Active		**Intransitive Middle**	
αἰσχύνω 'disgrace'	1	αἰσχύνομαι 'feel shame'	**3**
ἀπαλλάσσω 'set free from'	2	ἀπαλλάσσομαι 'depart'	**86**
αὐξάνω 'grow (trans.)'	1	αὐξάνομαι 'grow up'	**13**
ἐγείρω 'stir up'	2	ἐγείρομαι 'wake up, arise'	**5**
ἔλπω 'cause to hope/expect'	0	ἔλπομαι 'hope, expect'	**9**
ἵζω 'seat'	1	ἵζομαι 'sit down'	**39**
ἵστημι 'set, put'	47	ἵσταμαι 'stand'	**85**
κήδω 'trouble, distress'	0	κήδομαι 'care about'	**7**
κομίζω 'carry, fetch, take'	**40**	κομίζομαι 'travel'	26
λείπω 'leave behind'	**36**	λείπομαι 'stay behind'	16
λυπέω 'harass'	**6**	λυπέομαι 'be angry'	2
μιμνήσκω 'remind'	0	μιμνήσκομαι 'remember; mention'	**30**
ὁρμάω 'urge, dispatch'	2	ὁρμάομαι 'set off'	**93**
παύω 'stop'	14	παύομαι 'cease'	**45**
πείθω 'persuade, convince'	47	πείθομαι 'believe, obey'	**95**
πορεύω 'make to go, convey'	0	πορεύομαι 'travel, march'	**75**
σήπω 'make rot'	0	σήπομαι 'rot'	**4**
στέλλω 'send'	9	στέλλομαι 'set forth'	**20**
τήκω 'melt (trans.)'	2	τήκομαι 'melt (intr.)'	**3**
τρέπω 'turn (trans.)'	8	τρέπομαι 'turn (intr.)'	**94**
τρέφω 'grow (trans.), nurture'	**15**	τρέφομαι 'grow up'	11
φαίνω 'show'	14	φαίνομαι 'appear'	**141**
φοβέω 'alarm'	2	φοβέομαι 'fear'	**21**
φύω 'produce, grow (trans.)'	8	φύομαι 'grow (intr.)'	**28**
φέρω 'carry, bring'	**118**	φέρομαι 'fall, drift, charge'	16
ψεύδω 'deceive'	0	ψεύδομαι 'be mistaken'	**8**

In general, the middle intransitive forms occur far more frequently (printed in bold-face) in Herodotus than the active transitives. It strikes me as rather counter-intuitive to consider the intransitive middles as derived from, or somehow secondary to, the active transitive

[65] The point that both grammatically irregular forms (e.g. the English strong verbs) and regular forms are lexically stored is also made by Bybee (1985: 113-4, and *passim*).

[66] Note that most middle intransitive verbs have passive aorist forms in -(θ)η-. These forms are included in the frequency-figures.

counterparts, considering that intransitive middles generally have a higher token-frequency.[67] Although, in general, one could maintain that the middle voice is marked in comparison with the active voice (in terms of frequency, semantic content, and morphological markedness: see my discussion of markedness above (1.1.4)), in the case of these causative active and intransitive middle verbs the markedness-relation seems to be reversed. The frequency figures above clearly attest to that. Furthermore, the intransitive middles can be considered unmarked from a semantic point of view, as well. The event denoted by the intransitive middles (roughly: an autonomous change of state of the subject) is semantically more simple than the event denoted by the active causative verbs, since the latter add a causer-participant to the event. A third objection to the valence-reduction view on the middle voice is that many middle voice verbs are transitive, just as their active counterparts, like, for instance, the frequent and productive class of indirect reflexive middles. In conclusion, the valence-reduction view on the middle voice is unattractive. It is preferrable to seek a purely *semantic* explanation instead. In section 1.1.3, I argued that the middle voice can be explained adequately as a marker of subject-affectedness.[68]

[67] On the contrary, there are well-known cases in which an active transitive is created next to an old intransitive middle, e.g. ἥδω to ἥδομαι, μαίνω to μαίνομαι, πεύθω to πεύθομαι. These active causative vs. middle intransitive pairs contravene the global markedness of the middle voice. This phenomenon is can be called *local markedness* (see Tiersma 1982, Bybee 1985: 74-6).

[68] Sometimes, a more restricted view is adopted, namely that the *passive aorist* form exhibits valence reduction. This view is advocated by Andersen (1989, 1993). Yet, two of the objections that were discussed above are equally applicable. First, there are many media tantum with passive aorists (passiva tantum), e.g. ἐβουλήθην 'I wanted', ἐδυνήθην 'I could', ᾐδέσθην 'I was ashamed', ᾠήθην 'I thought'. Obviously, these passive aorists could never be derived from transitive counterparts. Secondly, the passive aorists often have a higher token-frequency than their active counterparts (e.g. ὥρμησα 'I urged, dispatched' vs. ὡρμήθην 'I started off'), which makes it unnatural to consider them as derivations of the active transitives. The third objection does not apply since passive aorists are not transitive. In chapter 3, I will argue that the passive aorist form expresses that the subject is similar to a prototypical patient.