

6 The learnability of grammatical and lexical stress in Modern Greek

6.1 Introduction

This chapter shows how underlying forms are learned by an OT on-line learning algorithm. The proposed algorithm is “stupid”: it processes one form at a time under one grammar at a time. “Stupid” is good and effective: no extra learning mechanisms are required than the ones already involved in a general grammar learning model of OT (e.g. Boersma 1997). Interpretation of incoming forms and constraint reranking as a reaction to error detection is enough. Surface and underlying forms are connected via faithfulness constraints, and underlying forms are connected to meaning via lexical constraints. It is shown that the learner can acquire grammar and underlying forms concurrently, and creates an economical lexicon. This is exemplified with the learning of lexical stress. Lexical stress (in opposition to grammatically assigned stress) is information that is not predictable by the grammar (i.e. the constraint ranking of a language), and is therefore stored in the lexicon as some sort of underlying representation.

In the learnability approaches of chapters 4 and 5 it was not necessary to stipulate a learning mechanism for underlying forms, since stress was always assigned by structural principles, and was not influenced by the lexicon. Underlying forms were always trivially derivable from the surface form and merely had to be disengaged from metrical structure in order to be looked up in the lexicon. If a language learner knew in advance whether she had to learn a language with grammatically assigned stress that would be all there is to it. If it were the case that stress was always determined by the lexicon, the learning of underlying forms would be trivial as well: the learner would simply have to map surface forms faithfully onto underlying forms. However, determining underlying forms is not always straightforward. First of all, the learner does not know whether her language has grammatical or lexical stress, and second, the language can have a mix of both. The learner needs to be able to handle this. A case in point is lexical stress in Modern Greek, where stress is not predictable by grammatical principles alone, but is highly influenced by the lexicon. Morphemes can contrast in being underlyingly stressed or unstressed, and can even be specified for being pre-

or post-stressing. Modern Greek stress therefore provides a challenging test ground for a learnability account of underlying forms; however, the model introduced here should be applicable to any other area where the learning of underlying forms is involved.

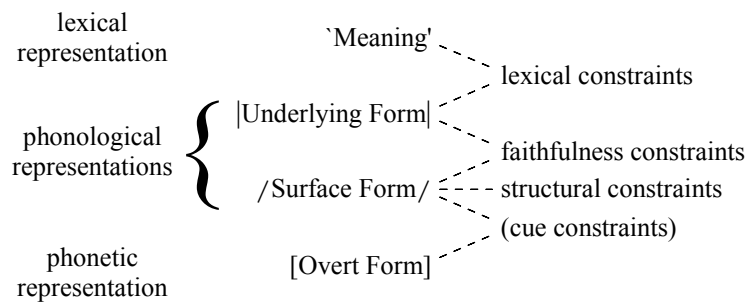
Proposals for the learning of underlying forms in OT are plenty, starting with e.g. Prince & Smolensky's *Lexicon Optimization* (1993; enhanced by Itô et al. 1995, and Tesar & Smolensky 2000), Tesar (2004, 2006), Tesar et al. (2003), or Jarosz (2006). A discussion of some of these alternatives is given in section 6.7. What they have in common is that they make use of paradigmatic information: the learner collects all the possible surface instantiations of a morpheme, and on the basis of the collected data adjusts the grammar (i.e. the constraint ranking). This kind of learning is called *off-line*: the gathered data are stored for further processing and can be accessed repeatedly in the course of learning. While the proposed algorithms for the learning of underlying forms are successful in what they are supposed to do, this kind of learning feels unnatural. When would a child know that she gathered enough information to be able to analyze the data? It is more natural to assume that a child processes a form the moment she is confronted with it, and discards the form immediately after processing. This is called *on-line* learning. By learning on-line, the child does not have to store any surface form. In the remainder of the chapter it is shown how such an on-line learning of underlying forms can be modelled in OT.

The proposed grammar model (as shown in figure (1) of chapter 1) is based on Boersma's grammar model (2005), but extended with an additional representational level 'Meaning', as also discussed in Boersma (2006ab). The idea is that we still have overt forms, surface forms, and underlying forms, but need an additional level of *meaning* to model the acquisition of underlying forms. Intuitively it means that you create a form only because you have a meaning that you want to express. Formally, the notion of meaning will help the learner to recognize identical morphemes, and reduce lexical allomorphy.

In this model, the different levels of representations are connected through different groups of constraints (shown in figure (171), which is repeated from figure (13) in chapter 2). In this chapter I will focus on all four levels of representation shown in (171): *meaning*, *underlying form*, *surface form* and *overt form*. The connection between overt form and surface form is expressed by *cue constraints* (Boersma 1998), which I ignore, because I model the relation between overt and surface forms with *structural*

constraints alone, as in chapters 4 and 5. The connection between surface and underlying forms is expressed by *faithfulness constraints*. The connection between form and meaning is expressed by a new family of lexical constraints (Boersma 2001). The term ‘meaning’ as used here conflates both semantic and syntactic information.

(171) The grammar model:



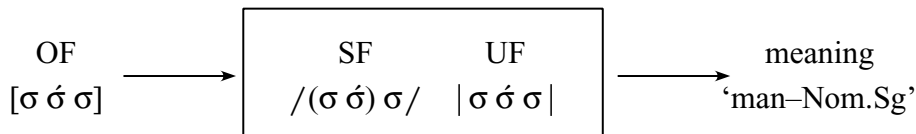
First, I provide an OT account of on-line learning of underlying forms where surface form and meaning are given, and the learner has to create the corresponding underlying form. The learner will deal with one level of hidden structure, as addressed in step 3 of the introduction and implemented in the learning approaches of surface forms in chapters 4 and 5. But instead of modelling hidden surface forms the learner will have to infer hidden underlying forms:

(172) Learning hidden underlying forms



As addressed in step 4 of the introduction and outlined in chapter 3, I push the proposal as far as to the learning of both surface and underlying form, given overt form and meaning. The learner then has to deal with two levels of hidden structure:

(173) Learning hidden surface forms and hidden underlying forms



The outline of the chapter is as follows: the next section provides a description and an analysis of lexical stress in Modern Greek. Section 6.3 outlines the learning model and its ingredients. The model is put to the test in computer simulations in sections 6.4, 6.5, and 6.6. Section 6.7 discusses some alternatives, and section 6.8 provides a general discussion.

6.2 Grammatical and lexical stress in Modern Greek

The learning of underlying forms is illustrated with a simplified version of Modern Greek stress. I will focus on the nominal paradigm, which will not weaken our claim made here, since nouns show all the stress patterns that are possible in Modern Greek. Modern Greek stress is weight-insensitive and vowels do not differ in phonological length.⁵⁷ Words in Modern Greek can be stressed on any of the last three syllables of a word (Joseph & Philippaki-Warburton 1987):

(174) A trisyllabic window for stress in Modern Greek

astráγalos ‘ankle–Nom.Sg.M’

ániksi ‘spring–Nom.Sg.F’

mitéra ‘mother–Nom.Sg.F’

mixaní ‘machine–Nom.Sg.F’

⁵⁷ They might differ in phonetic duration, though. Stressed syllables tend to be phonetically longer than unstressed ones. In fact, stress in Modern Greek is mainly phonetically realized as a combination of duration and intensity (McKeever Dauer 1980; Botinis 1989). Nevertheless, there is no phonemic vowel length distinction.

Some words keep their stress throughout the paradigm:

(175) Fixed stress

γόνδολα ‘gondola-Nom.Sg’ vs. *γόνδολον* ‘gondola-Gen.Pl’

Others shift stress:

(176) Stress shift

θάλασα ‘sea-Nom.Sg’ vs. *θαλασόνα* ‘sea-Gen.Pl’

The common view on Modern Greek is that stress is mainly assigned through specifications in the lexicon (e.g. Philippaki-Warbuton 1970, 1976, Ralli 1988, Malikouti-Drachman & Drachman 1989, Touratzidis & Ralli 1992, Drachman & Malikouti-Drachman 1996, Revithiadou 1999). Roots as well as suffixes can contrast in stress. I follow Revithiadou (1999) in the classification that roots and suffixes can be unstressed or stressed. For computational reasons I simplify the specifications for morphemes that happen to push stress onto another morpheme: roots that are “unaccentable” under Revithiadou’s analysis will simply be specified as post-stressing, and suffixes that are specified for having a “weak accent” in her analysis will be specified for being pre-stressing. In Revithiadou’s account, structural constraints are responsible for unaccentable roots; in my account given here, these roots are subject to faithfulness constraints.

I first focus on the contrast between underlyingly stressed and unstressed morphemes. A word like *γόνδολα* ‘gondola-Nom.Sg’ retains stress on the root when inflected with the genitive plural suffix *-on*: *γόνδολον*, as in example (175). But a word like *θάλασα* ‘sea-Nom.Sg’ shifts stress to the suffix when inflected: *θαλασόνα*, as in (176). The root of a word like *γόνδολα* is analyzed as being underlyingly stressed [*γόνδολ-*], and the root in a word like *θάλασα* is analyzed as being underlyingly unstressed [*θάλας-*]. The genitive plural suffix *-on* is underlyingly stressed, as becomes apparent when attached to an unstressed root like in the case of *θαλασόνα*: only then can it surface as stressed. The nominative singular suffix *-a* is underlyingly unstressed, as becomes apparent when combined with an unstressed root *θάλας-*: then the phonological default stress on the antepenultimate syllable is assigned. In the case that both root and suffix are specified for stress, as in *γόνδολον*, the root stress is preserved.

I will ignore foot structure for the moment (but will come back to it in section 6.6), and represent surface forms (SF) as in column a. of table (177). Underlying forms (UF) are represented as in column b. of table (177), and meaning as in column c. of table (177). The notion ‘meaning’ here refers thus to both the semantic content of a root (e.g. *ɣondol-* expresses the concept ‘gondola’) and to the syntactic content of a suffix (e.g. *-on* expresses case ‘genitive’ and number ‘singular’).

(177) A simple contrast⁵⁸

a. surface forms	b. underlying forms	c. meaning
/ɣóndola/	ɣóndol+a	‘gondola-Nom.Sg’
/ɣóndolon/	ɣóndol+ón	‘gondola-Gen.Pl’
/θálasa/	θalas+a	‘sea-Nom.Sg’
/θalason/	θalas+ón	‘sea-Gen.Pl’

We can distil four underlying morphemes from the table in (177): an underlyingly stressed root |ɣóndol-|⁵⁹, an unstressed root |θalas-|, a stressed suffix |-ón|, and an unstressed suffix |-a|. The learner is expected to eventually arrive at the four proposed morphemes, but cannot know in advance whether e.g. the root *ɣondol-* is specified for stress or not. The two-way contrast in morphemes means for the learner that she can choose between a stressed |ɣóndol-| and an unstressed |ɣondol-|. The same applies to every other morpheme, which gives our learner a pool of eight possible underlying forms to choose from:

(178) Possible underlying forms:

ɣondol-	θalas-	-a	-on
ɣóndol-	θálas-	-á	-ón

My goal is to establish only a basic analysis for Modern Greek stress that can easily be implemented into the computer simulations, because I want to focus on the learning of underlying forms, and not on a detailed analysis of Modern Greek stress. The reader is kindly asked to forgive me if I therefore

⁵⁸ The contrast described in this section is nothing else than the language in example (5) of the PAKA-world in Tesar et al. (2003:480).

⁵⁹ If I were to model the acquisition of stops in Modern Greek, I would analyze the coronal stop /d/ as underlyingly voiceless (Arvaniti 1999). Since this is beyond the scope of the dissertation, I stick to the voiced surface form.

leave out issues such as secondary stress (which is controversial in Modern Greek) and the influence of derivational suffixes (which draw stress away from roots; see Revithiadou 1999 for a detailed analysis, and Apoussidou 2003 for an alternative account). Section 6.2.1 provides an analysis of the phonological default stress. Section 6.2.2 provides an analysis of lexically stressed words.

6.2.1 Analyzing grammatical stress in Modern Greek

The default stress on the antepenultimate syllable (or on the penultimate syllable if the word consists of two syllables) applies when none of the morphemes of a word are underlyingly stressed. To account for the limitation of stress to the last three syllables of a word, and to make sure that stress will be on the antepenultimate syllable in a word without lexical stress, we need the following universal structural constraints (which is a limited choice from the constraints I used in chapters 4 and 5):

- (179) Structural constraints for Modern Greek
 AFL/AFR: Feet are aligned at the left/right word edge
 FTBIN: Feet are disyllabic.
 IAMBIC: Feet are right-headed.
 NONFINAL: No foot is final in a prosodic word.
 TROCHAIC: Feet are left-headed.

AFR needs to be ranked above AFL, because stress is not left-aligned:

- (180) AFR >> AFL

a.stra.ɣal+os	AFR	AFL
☞ /a (strá.ɣa) los/	*	*
/((á.stra) ɣa.los/	**!	

NONFINAL has to outrank AFR, because the final syllable is extra-metrical:

- (181) NONFINAL >> AFR

θa.las+a	NONFINAL	AFR
☞ /((θá.la) sa/		*
/θa (lá.sa)/	*!	

TROCHAIC outranks IAMBIC, because feet are left-headed:

(182) TROCHAIC >> IAMBIC

θa.las+a	TROCHAIC	IAMBIC
☞ / (θá.la) sa/		*
/ (θa.lá) sa/	*!	

FTBIN outranks TROCHAIC, to rule out monosyllabic feet:

(183) FTBIN outranks TROCHAIC

θa.las+a	FTBIN	TROCHAIC
☞ / (θá.la) sa/		
/ (θá) la.sa/	*!	

If we wrap it up in one tableau, it still works:

(184) NONFINAL >> AFR; FTBIN >> TROCHAIC >> IAMBIC

θa.las+a	NONFIN	AFR	AFL	FTBIN	TROCH	IAMBIC
☞ / (θá.la) sa/		*				*
/ θa (lás.a) /	*!		*			*
/ θa.la (sá) /	*!		**	*		
/ (θa.lá) sa /		*			*!	
/ θa (la.sá) /	*!		*		*	
/ (θá) la.sa /		**!		*		

6.2.2 Analyzing lexical stress in Modern Greek

If any morpheme in a word has lexical (= underlying) stress, surfacing stress is determined by faithfulness, which outranks the structural constraints. Or rather, MAX(Stress) outranks the structural constraints; DEP(Stress) can be violated (as we can see from the default case of / (θá.la) sa/, where an accent is inserted on the root morpheme) and has to be ranked low. The high ranking of MAX(Stress) is illustrated in tableau (185), where I only include those structural constraints in the tableau that are high-ranked and violated. The first candidate is ruled out because it violates MAX(Stress): the underlying stress specification of the suffix is not realized on the suffix in

the output form. The winning candidate violates a number of structural constraints because it puts a monosyllabic foot into final position, but it satisfies the highest-ranked MAX(Stress) constraint.

(185) MAX(Stress) outranks the structural constraints

θa.las+ón	MAX(Stress)	NONFIN	FTBIN
/(θá.la) son/	*!		
☞ /θa.la (són)/		*	*

When a lexically stressed root and a lexically stressed suffix are combined, only one of them can surface. The root retains its stress, as we can see from *ýndolon* in tableau (186). Therefore, the MAX constraint needs to be split up into MAX(Root) and MAX(Affix). Likewise, DEP is split up into DEP(Root) and DEP(Affix). The DEP constraints are ranked below the MAX constraints.

(186) MAX(Root) >> MAX(Affix)

ýón.doI+ón	MAX(Root)	MAX(Affix)
☞ /(ýón.do) lon/		*
/ýon.do (lón)/	*!	

The next section describes the proposed model for the learning of underlying forms.

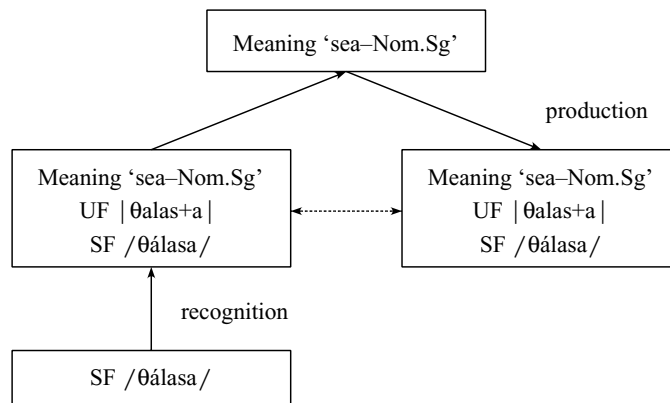
6.3 The model for the learning of underlying forms

In the proposed model the learner learns the underlying forms of a language by ranking the relevant constraints.

For the purposes of 6.4 and 6.5, I ignore foot structure. Thereby the surface forms (henceforth ‘SF’) as in column a. of table (177) are observable to the learner. The mapping between SFs and underlying forms (‘UF’s as in column b. of table (177)) is regulated by the faithfulness constraints listed in (16) of section 2.6.2, and the mapping between UFs and meaning (meaning as in column c. of table (177) above) is regulated by the lexical constraints discussed in 2.6.3. The learning process is error-driven like described in section 3.5.2, in the sense that a learner only changes her grammar if she detects a mismatch between the form she recognizes and the form she would

produce for that word. In recognition, the SF serves as input to the OT evaluation, and the candidates are triplets of meaning, UF, and SF. In virtual production, i.e. the computation of the form that the learner would have produced herself, meaning serves as input to the evaluation, and the candidates are the same kind of triplets as in recognition. If the meaning/UF/SF triplet of the virtual production step is identical to the one in recognition, nothing is changed in the grammar. If there is a mismatch (i.e. an error is detected) the grammar (i.e. the constraint ranking) is changed.

(187) The processing model



I propose that at least two kinds of constraints are involved in the learning of underlying forms: faithfulness and lexical constraints. Faithfulness constraints are necessary to establish a correspondence between SF and UF, in this case faithfulness constraints on stress. The faithfulness constraints need to be split up into MAX and DEP since deletion of stress in the surface form (incurring a violation of MAX) as well as insertion of stress can occur (incurring a violation of DEP). MAX and DEP have to be furthermore split up into faithfulness for roots and faithfulness for affixes, as shown in tableau (186); in my simplified version of Modern Greek, faithfulness to the root outranks faithfulness to the affix. We have already seen these constraints in section 2.6.2. The following section goes into detail about lexical constraints.

6.3.1 Lexical constraints

Lexical constraints establish the link between UF and meaning. Boersma's (2001) approach to the phonology-semantics interaction in OT makes use of lexical constraints to distinguish between two homophonous forms with different meanings in comprehension, such as Dutch [rat] 'wheel' and [rat] 'rat'. The overt form [rat] 'wheel' has a voiced coda underlyingly, [rad], while the overt form [rat] 'rat' is underlyingly [rat]. If both forms equally violate any structural constraints, and if the decision between the two forms is solely left to faithfulness constraints in comprehension, the form [rat] 'wheel' will never be chosen, because it incurs a faithfulness violation to the coda consonant, which is underlyingly voiced. The choice would always fall on [rat] 'rat', because this form does not incur a faithfulness violation. Boersma therefore proposed lexical access constraints militating against each possible underlying form in order to model the access of either [rat] 'wheel' or [rat] 'rat' depending on the ranking of these constraints and the semantic context in which the two forms are encountered.

Escudero (2005:220ff.) similarly modelled *message-driven learning in recognition* involving faithfulness constraints and lexical constraints. She proposed that the perception grammar of a (for instance, Dutch) speaker learning a new language (for instance, Spanish) can be adjusted by the influence of the semantic context in which a given form occurs. Initially, the second-language learner brings her native grammar and lexicon and applies them to the second language. In Escudero's account, learners of a new language can adjust their (already present) phonological categories with the help of semantic context. Learners learn from comprehension alone; the recognition process influences the perception process.

The situations in Boersma's (2001) and Escudero's (2005) approaches to the determination of underlying forms differ from the L1 learning situation assumed here. A child learning her first language does not initially have any phonological categories, as little as she has underlying forms. She is not (only) confronted with the problem of homonymy, but with the problem of determining underlying forms in the first place. Given the overt form that already includes stress, her recognition grammar cannot have any influence on her perception.⁶⁰ I therefore assume constraints against the

⁶⁰ In the case of acquiring hidden surface forms in a second language, Boersma's or Escudero's approach will not help, either: lexical stress is phonetically not distinguishable from grammatical stress. Their proposals work if one models the

creation of underlying forms that will enable the learner to establish the appropriate underlying forms for her language. The lexical constraints I employ are formulated as follows:

- (188) Lexical constraints on underlying stress specifications:
Don't connect the meaning 'xy' to the form |XY| that is specified/unspecified for stress.

For each meaning (or morpheme, for that matter) several constraints are induced that militate against specified corresponding underlying forms. In the case at hand, the constraints militate against the connection of a meaning to an underlying form that is specified for being stressed or unstressed. For a form like *thalasón*, there are constraints on roots and affixes. Two constraints on the relation between a meaning and a root morpheme are given in (189):

- (189) Lexical constraints on the root *thalas-*:
 *|*thalas-*| 'sea': *Don't connect the meaning 'sea' to an unstressed root |thalas-|.*
 *|*thálas-*| 'sea': *Don't connect the meaning 'sea' to a stressed root |thálas -|.*

Likewise, there are several constraints on the affix *-on*:

- (190) Lexical constraints on the suffix *-on*:
 |-on*| 'Gen.Pl': *Don't connect the meaning 'Gen.Pl' to an unstressed suffix |-on|.*
 |-ón*| 'Gen.Pl': *Don't connect the meaning 'Gen.Pl' to a stressed suffix |-ón|.*

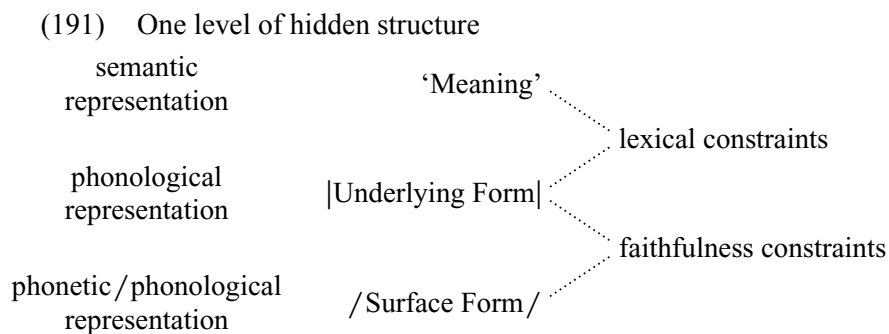
As already mentioned in section 2.6.3, these constraints are not necessarily innate, but could be induced by the learner. Whenever a learner encounters a new meaning and a new form, she invokes a constraint against the connection between the two. For this she needs to be equipped with the ability to invoke constraints against any possible underlying form. In lack of

acquisition of the phonetic correlates for stress, or in the case of second language acquisition, when the phonetic correlates of the second language differ from the ones in the first language; then message-driven recognition learning can help to adjust the correlates for stress.

an algorithm that does so, the required constraints are implemented in the simulations.

6.3.2 Recognition

I first focus on the mapping of the surface form onto an underlying form and meaning here. This implies that overt forms are set aside for the moment. Figure (191) illustrates this part of the bigger model given in figure (171).



The recognition mapping proceeds as follows: the listener interprets an incoming SF as an UF by applying her grammar.⁶¹ This means that the SF becomes the input to an OT evaluation, with triplets of meaning/UF/SF as candidates (provided by GEN). A ranking of faithfulness constraints in interaction with lexical constraints selects the optimal meaning/UF/SF triplet. In (192), the evaluation of the incoming SF /θalasón/ is shown. The candidate meaning/UF/SF triplets only differ in their UFs. The UFs are split up into roots and affixes, and since there is a two-way contrast both in roots and in affixes (stressed/unstressed), there are four possible candidates that have the same surface form (because this is the input) and the same meaning (because I ignore the possibility of homonymy).⁶² Consider a listener with

⁶¹ This is in fact the same function as Robust Interpretive Parsing or the perception process in the mapping of overt forms to surface forms that we have seen in sections 3.2, 4.5, and 5.3.1, except that this time, the input is the surface form, and the form to be interpreted is the underlying form.

⁶² I simplifyingly assume at this point that only the initial syllable of the root can be specified for stress. In real Modern Greek, however, any of the syllables in disyllabic roots can be specified for stress.

the constraint ranking as in tableau (192). All faithfulness constraints are ranked high. This rules out any candidate that is not faithful to the underlying form: the second and the last candidate are excluded by high-ranked MAX(Root), because the underlying stress on the root is not realized in the surface form; the first candidate is ruled out by high-ranked DEP(Affix), because the surface form has stress on the affix, but the underlying affix is unstressed. The listener recognizes the third candidate with an underlyingly unstressed root and an underlyingly stressed affix, [θalas+ón].

(192) Recognition of *thalasón*

SF /θalasón/	MAX(R)	MAX(A)	DEP(R)	DEP(A)	* θalas- 'sea'	* θálas- 'sea'	* -on 'Gen.PI'	* -ón 'Gen.PI'
'sea-Gen.PI' θalas+on /θalasón/				*!	*		*	
'sea-Gen.PI' θálas+on /θalasón/	*!			*		*	*	
☞ 'sea-Gen.PI' θalas+ón /θalasón/					*			*
'sea-Gen.PI' θálas+ón /θalasón/	*!					*		*

An adult listener will ideally recognize the same form that she would produce. A learner is already a listener, too, and will proceed the same way in recognition as an adult speaker. However, as outlined in sections 3.3 to 3.5, a learner has an under-developed grammar and might compute a different form in the production step. This is demonstrated in the next section.

6.3.3 Virtual production

As discussed in chapter 3, a language learner will run a check on her grammar as soon as she recognized a word by virtually producing it. The meaning becomes the input to the production evaluation. Again, the learner can choose between meaning/UF/SF triplets by applying her grammar (the same constraint ranking as in the recognition step). The upper left cell, the place reserved for the input to an evaluation, contains now 'meaning', based

on the idea that if you intend to produce an utterance, you start out with the meaning you want to express, and hence choose a form to say it with. GEN provides all possible combinations of meaning, underlying forms and surface forms as candidates. The underlying forms are combinations of root and suffix, and differ with respect to their stress specification. Unstressed roots are combined with unstressed or stressed suffixes, as are stressed roots. There are two different surface forms, one realizing stress on the root and one realizing stress on the suffix. I do not model structural constraints on the surface forms (yet). Since for the moment I only want to model a two-way contrast in morphemes, I exclude the possibility of stress on the second syllable of the root; I simply bar this possibility from GEN. There are more candidates in the production evaluation than in the recognition evaluation because in production both the UFs and the SFs can vary. If the optimal meaning/UF/SF triplet is identical to the one in recognition, no reranking takes place. If the two triplets are not identical, the resulting mismatch is deemed an error (as outlined in section 3.5), and the learner will adjust her constraint ranking. This is shown in tableau (193): the candidate with an underlyingly stressed root and an underlyingly unstressed suffix is the winner of the virtual production step, (indicated by ‘ \ominus ’), and differs from the winner in recognition (indicated by ‘ \wp ’).

(193) Production of ‘sea-GEN.PL’

‘sea-Gen.PL’	MAX(R)	MAX(A)	DEP(R)	DEP(A)	* θ alas- ’sea’	* θ alas- ’sea’	* -on ’Gen.PL’	* -ón ’Gen.PL’
‘sea-Gen.PL’ θ alas+on / θ alasón/				*!	*		*	
‘sea-Gen.PL’ θ alas+on / θ álason/			*!		*		*	
‘sea-Gen.PL’ θ álas+on / θ alasón/	*!			*		*	*	
\ominus ‘sea-Gen.PL’ θ álas+on / θ álason/						←*	←*	
\wp ‘sea-Gen.PL’ θ alas+ón / θ alasón/					*!→			*→
‘sea-Gen.PL’ θ alas+ón / θ álason/		*!	*		*			*
‘sea-Gen.PL’ θ álas+ón / θ alasón/	*!					*		*
‘sea-Gen.PL’ θ álas+ón / θ álason/		*!				*		*

This brings about constraint reranking. Any difference in the triplets of the recognized candidate from the virtually produced candidate (i.e., if the UF

that is chosen in production differs from the UF in recognition, *or* the SF chosen in production differs from the SF in recognition) elicits error detection and subsequent constraint reranking. The reranking makes it more likely that in the future, recognition and production are brought into harmony.

The recognition/production step that models three levels of representation, i.e. surface form, underlying form and meaning, is put to the test in the next section.

6.4 Testing a two-way contrast

For the purpose of testing a two-way contrast in stress, I assume that the learner only has to find out whether a morpheme is underlyingly stressed or not. Therefore only faithfulness and lexical constraints are included in the simulations, and no structural constraints. I put 10 virtual learners (created in the Praat programme; Boersma & Weenink 1992-2006) to the test. The virtual learners reranked the constraints in a GLA-fashion (Boersma 1997) in the sense of section 3.6.2, with the slight difference that constraints were reranked with weighted uncancelled.⁶³ All constraints were initially ranked at the same heights, and the learners learned from tableaux as shown in (192) and (193). Learning was done with an evaluation noise of 2.0 and an initial plasticity of 1.0 with a decrement of 0.1. The group of learners was homogeneous, i.e. there was no variation in the reranking strategy or constraint set.

6.4.1 The training data

All of the virtual GLA learners had access to the same training data consisting of four different surface forms listed in (194), which we have

⁶³ ‘Weighted uncancelled’ means that the ranking is lowered for all constraints that are violated more in the recognized form than in the learner’s production, and the ranking is raised for all the constraints that are violated more in the learner’s production than in the recognized form, and the size of the learning step is divided by the number of constraints that move in the same direction. This makes sure that the average ranking of all the constraints is constant.

already encountered in section 6.2. The learners encountered the data in a different order. For each learner a total of 1 000 000 forms was drawn randomly from the training set, i.e. each of the four forms was presented approximately 250 000 times.

(194) The training data:

/yóndola/ /yóndolon/ /θálasa/ /θalason/

The learners were furthermore equipped with the knowledge that corresponding underlying forms are composed of roots and suffixes, and with the meaning of these morphemes. The possible underlying forms that the learners could create are provided in the next section.

6.4.2 The pool of underlying forms to choose from: GEN

For testing the two-way contrast, GEN contained eight possible underlying forms that the learners could choose from, listed in (195). Each root and each suffix exists as underlyingly stressed or unstressed.

(195) Possible underlying forms:

yondol-	θalas-	-a	-on
yóndol-	θálas-	-á	-ón

Simplifyingly I assume that GEN contains only candidates where roots can have stress on the first syllable, not the second. The following section provides the constraint set.

6.4.3 The constraint set

The learners of the two-way contrast are equipped with a constraint set including faithfulness and lexical constraints. The faithfulness constraints are listed in (196), and have been discussed in section 2.6.2.

(196) The faithfulness constraints on stress

MAX(Root)	DEP(Root)
MAX(Affix)	DEP(Affix)

The lexical constraints have been discussed in section 2.6.3. The full list for the simulations of the two-way contrast is given in (197).

- (197) Lexical constraints:
- *|γondol-| ‘gondola’: Don’t connect the meaning ‘gondola’ to an unstressed root |γondol-|.
 - *|γóndol-| ‘gondola’: Don’t connect the meaning ‘gondola’ to a stressed root |γóndol-|.
 - *|θalas-| ‘sea’: Don’t connect the meaning ‘sea’ to an unstressed root |θalas-|.
 - *|θálas-| ‘sea’: Don’t connect the meaning ‘sea’ to a stressed root |θálas-|.
 - *|-a| ‘Nom.Sg’: Don’t connect the meaning ‘Nom.Sg’ to an unstressed suffix |-a|.
 - *|-á| ‘Nom.Sg’: Don’t connect the meaning ‘Nom.Sg’ to a stressed suffix |-á|.
 - *|-on| ‘Gen.Pl’: Don’t connect the meaning ‘Gen.Pl’ to an unstressed suffix |-on|.
 - *|-ón| ‘Gen.Pl’: Don’t connect the meaning ‘Gen.Pl’ to a stressed suffix |-ón|.

The list should in principle also contain constraints such as *|γondol-| ‘sea’ (‘don’t connect a meaning ‘sea’ to the root |γondol-|’). I saved some ink and computation time by not including these constraints (and they would always end up top-ranked in the computer simulations anyway, because the forms are always given along with their meaning).

6.4.4 Results: the chosen underlying forms

The final ranking of one example learner is shown in (198) (the constraints at the top are ranked higher than the constraints at the bottom):

(198) Ranking learner No. 1:

- *|γondol-| ‘gondola’
- *|-á| ‘Nom.Sg’
- MAX(R)
- DEP(A)
- *|-on| ‘Gen.Pl’
- *|θálas-| ‘sea’
- *|-a| ‘Nom.Sg’
- *|γóndol-| ‘gondola’
- DEP(R)
- MAX(A)
- *|-ón| ‘Gen.Pl’
- *|θalas-| ‘sea’

If we evaluate e.g. *γóndolon* with this ranking in tableau (199), we can see that the candidate with an underlyingly stressed root and an underlyingly stressed affix is chosen, together with the correct surface form (I leave out meaning in the candidate cells because it is always the same anyway). The numbers above the constraint columns indicate the final ranking values of the constraints. Lexical constraints militating against other morphemes are left out for better readability. All candidates with an underlyingly unstressed root are ruled out by high-ranking *|γondol-| ‘gondola’, irrespective of their surface form. The next-ranking constraint MAX(R) rules out the candidates that have an underlyingly stressed root, but where the root stress is not retained in the surface form. That leaves two competing candidates with identical surface stress on the root: one where the root is underlyingly stressed and the suffix is unstressed, and one where both root and suffix are underlyingly stressed. The constraint against an unstressed suffix *|-on| ‘Gen.Pl’ is ranked higher than the constraints against a stressed suffix *|-ón| ‘Gen.Pl’. This results in the winning candidate |γóndol+ón| /γóndolon/, where both morphemes are underlyingly stressed and where stress is on the root in the surface form.

(199) A meaning-to-form evaluation

	105.655	105.597	105.571	103.153	96.454	95.721	95.711	93.523
‘gondola-Gen.Pl’	* yondol- ‘gondola’	MAX(R)	DEF(A)	* -on ‘Gen.Pl’	* yondol- ‘gondola’	DEF(R)	MAX(A)	* -ón ‘Gen.Pl’
yondol+on /yóndolon/	*!			*		*		
yondol+on /yondolón/	*!		*	*				
yóndol+on /yóndolon/				*!	*			
yóndol+on /yondolón/		*!	*	*	*			
yondol+ón /yóndolon/	*!					*	*	*
yondol+ón /yondolón/	*!							*
☞ yóndol+ón /yóndolon/					*		*	*
yóndol+ón /yondolón/		*!			*			*

I checked how stably this learner would choose her candidates. Out of 1 000 trials with an evaluation noise of 2.0, the learner virtually produced in 99.9% of the cases |yóndol+ón| /yóndolon/ for the meaning ‘gondola-Gen.Pl’. In 0.1% of the cases, however, MAX(A) outranked *|-on| ‘Gen.Pl’, and she therefore chose for the competing form |yóndol+on| /yóndolon/, where the surface form is the same, but the underlying form consists of a stressed root and an unstressed suffix. I regard this as a licit slip of the mind. The overall percentages for the forms that this learner chose are given in (200):

(200) Percentages for learner No. 1

‘gondola-Nom.Sg’	→ yóndol+a /yóndola/	100%
‘gondola-Gen.Pl’	→ yóndol+ón /yóndolon/	99.9%
‘sea-Nom.Sg’	→ θálas+a /θálasa/	99.5%
‘sea-Gen.Pl’	→ θálas+ón /θálasón/	99.9%

All the learners created grammars with the crucial rankings displayed in (201), and correct percentages similar to learner No. 1. Note that it is not important whether a constraint like *|yondol-| ‘gondola’ is ranked above a constraint *|θálas-| ‘sea’, but only whether they are ranked with another constraint referring to the same meaning.

(201) Crucial rankings

- *|yondol-| ‘gondola’ >> *|yóndol-| ‘gondola’
 *|θálas-| ‘sea’ >> *|θálas-| ‘sea’
 *|-á| ‘Nom.Sg’ >> *|-a| ‘Nom.Sg’
 *|-on| ‘Gen.Pl’ >> *|-ón| ‘Gen.Pl’
 MAX(R) >> MAX(A)
 DEP(A) >> DEP(R)

Since the learners choose to almost a 100% of the times the underlying forms as established (by linguists) in section 6.2, I find this an encouraging result and will check what happens when there is a three-way contrast in underlying forms involved: stressed/unstressed and pre-/post-stressing forms.⁶⁴ This is a step closer to real Modern Greek and is shown in the next section.

6.5 Testing a three-way contrast: learning pre- and post-stressing morphemes

As already mentioned in section 6.2, there are morphemes in Modern Greek that push stress onto another morpheme, instead of being stressed themselves. Examples of post-stressing roots are *uranós* ‘sky-Nom.Sg’ and *αγορά* ‘market-Nom.Sg’. The suffixes *-os* and *-a* behave as underlyingly unstressed in combination with underlyingly unstressed roots such as *άνθρωπος* ‘man-Nom.Sg’ and *θάλασα* ‘sea-Nom.Sg’, therefore it is unlikely (but not completely unreasonable to assume) that the suffixes in *uranós* and *αγορά* are underlyingly stressed allomorphs of the underlyingly unstressed suffixes in *άνθρωπος* and *θάλασα*. Rather, the stress pattern of these words is a property of the roots. I analyze pre- and post-stressing morphemes in the same way as I analyze underlyingly stressed morphemes: The root is lexically specified for being post-stressing, indicated with ‘→’, and MAX is responsible for the surfacing stress pattern. The winning candidate in (202) is the one with stress on the suffix.

⁶⁴ This looks like a four-way contrast, but is not really: I assume post-stressing underlying forms only for roots, and pre-stressing underlying forms only for suffixes.

(202) Post-stressing roots

a.ɣor→+a	MAX(R)	NONFINAL	FTBIN	DEP(A)
/ (á.ɣor) a/	*!			
☞ /a.ɣor (á)/		*	*	*

Likewise with pre-stressing suffixes: they impose a stress on the preceding syllable with the help of high-ranking MAX(A), as illustrated in tableau (203). Pre-stressing suffixes are indicated by a '←'. MAX(A) is violated whenever the underlying stress specification of the suffix is not fulfilled in the surface form; for tableau (203) it means that the constraint is violated when stress is too far to the left in the surface form, as in the first candidate, and also when the stress is realized on the suffix itself, as in the second candidate.⁶⁵

(203) Pre-stressing suffixes

an.θrop+←-u	MAX(A)	NONFINAL	FTBIN	DEP(A)
/ (án.θro) pu	*!			
/an.θro (pú)/	*!	*	*	*
☞ /an (θró.pu)/		*		

For the purpose of testing the three-way contrast in stress, I assume that the learners have to find out whether a root is underlyingly post-stressing, stressed, or unstressed and whether a suffix is underlyingly pre-stressing, stressed, or unstressed. Again, only faithfulness and lexical constraints are included in the simulations, and no structural constraints. I put another 10 virtual GLA learners (created in the Praat programme; Boersma & Weenink 1992-2006) to the test. All constraints were initially ranked at equal heights, and the learners learned from tableaux introduced in (192) and (193). Learning was done with an evaluation noise of 2.0 and an initial plasticity of 1.0 with a decrement of 0.1. The group of learners was homogeneous, i.e. there was no variation in the reranking strategy or constraint set.

⁶⁵ This means here that MAX, together with the specification, is also responsible for the exact place of the stress within the word.

6.5.1 The training data

The training data consisted of seven surface forms that included stress, but excluded foot structure. All learners encountered the data in different orders. For each learner a total of 1 000 000 forms was drawn randomly from the training set, i.e. each of the seven forms was presented approximately 140 000 times.

- (204) The training data⁶⁶
- | | | | |
|----------|---------|----------|-------|
| γόνδολα | θάλασα | άνθρωπος | αγορά |
| γόνδολον | θαλασόη | ανθρώπου | |

The learners were again equipped with the knowledge that the underlying forms are composed of roots and suffixes, and with the meaning of these morphemes. The possible underlying forms that the learners could create are provided in the next section.

6.5.2 The pool of underlying forms to choose from: GEN

Our regular virtual learner in this section does not know whether the root of *γόνδολα* is underlyingly stressed, unstressed or post-stressing, but she knows that it *could* be stressed, unstressed or post-stressing. I add post-stressing versions of all the roots in (195), e.g. |αγορ→|, which combines with unstressed |-a|, and a pre-stressing version for every suffix in (195), e.g. |←-on|. I also add a suffix |-u| with the meaning ‘Gen.Sg.’ and a root |ανθροπ-|, meaning ‘human’, to my universe. For each root, GEN therefore provides three possible allomorphs either stressed, unstressed, or post-stressing. For suffixes, GEN provides three possible allomorphs either stressed, unstressed, or pre-stressing. The list is given in (205).

⁶⁶ The feminine nouns ending in *-a* do not take the pre-stressing suffix *-u* as the genitive singular form (they take *-as*). Masculine nouns in *-os* take *-on* as the genitive plural suffix, but show different stress behaviour, e.g. *άνθρωπος* becomes *ανθρώπον*. It seems that there are indeed allomorphs for the suffix *-on*: an underlyingly stressed one *-ón*, and underlyingly pre-stressing one, *←-on*. Since the roots cannot be freely combined with any of the suffixes, only a subset of all theoretically thinkable combinations is included in the simulations.

(205) Underlying forms as of ‘freedom of analysis’

ɣondol-	θalas-	-a	-on
ɣóndol-	θálas-	-á	-ón
ɣondol→	θalas→	←a	←on
aɣor-	anθrop-	-u	-os
áɣor-	ánθrop-	-ú	-ós
aɣor→	anθrop→	←u	←os

For each possible underlying form there are lexical constraints that militate against it. The relevant ones are listed in the next section.

6.5.3 The constraint set

The learners are again equipped with the faithfulness constraints listed in (196) and the lexical constraints listed in (197). Since there is a three-way contrast this time, more lexical constraints have to be added. The relevant ones are listed in (206). This results in a set of 24 lexical constraints in total, plus four faithfulness constraints (in total 28 constraints). The learners learned from tableaux as shown in sections 3.3 to 3.5.

(206) Additional lexical constraints

- *|ɣondol→| ‘gondola’: Don’t connect the meaning ‘gondola’ to a post-stressing root |ɣondol→|.
- *|θalas→| ‘sea’: Don’t connect the meaning ‘sea’ to a post-stressing root |θalas→|.
- *|aɣor-| ‘market’: Don’t connect the meaning ‘market’ to an unstressed root |aɣor-|.
- *|áɣor-| ‘market’: Don’t connect the meaning ‘market’ to a stressed root |áɣor-|.
- *|aɣor→| ‘market’: Don’t connect the meaning ‘market’ to a post-stressing root |aɣor→|.
- *|anθrop-| ‘human’: Don’t connect the meaning ‘human’ to an unstressed root |anθrop-|.
- *|ánθrop-| ‘human’: Don’t connect the meaning ‘human’ to a stressed root |ánθrop-|.
- *|anθrop→| ‘human’: Don’t connect the meaning ‘human’ to a post-stressing root |anθrop→|.
- *|←a| ‘Nom.Sg’: Don’t connect the meaning ‘Nom.Sg’ to a pre-stressing suffix |←a|.
- *|←on| ‘Gen.Pl’: Don’t connect the meaning ‘Gen.Pl’ to a pre-stressing suffix |←on|.
- *|-os| ‘Nom.Sg.M’: Don’t connect the meaning ‘Nom.Sg.M’ to an unstressed suffix |-os|.
- *|-ós| ‘Nom.Sg.M’: Don’t connect the meaning ‘Nom.Sg.M’ to a stressed suffix |-ós|.
- *|←os| ‘Nom.Sg.M’: Don’t connect the meaning ‘Nom.Sg.M’ to a pre-stressing suffix |←os|.
- *|-u| ‘Gen.Sg’: Don’t connect the meaning ‘Gen.Sg’ to an unstressed suffix *|-u|.
- *|-ú| ‘Gen.Sg’: Don’t connect the meaning ‘Gen.Sg’ to a stressed suffix |-ú|.
- *|←u| ‘Gen.Sg’: Don’t connect the meaning ‘Gen.Sg’ to a pre-stressing suffix |←u|.

6.5.4 Results: the chosen underlying forms

When we check the outputs of the 10 different learners, we can see that the learners acquired the correct surface forms, but often chose between two different underlying forms for a meaning or between two surface forms for a meaning. This is illustrated with learner No. 1's ranking, listed in (207)⁶⁷ and evaluated in tableau (208).

(207) The ranking values for learner No. 1

* γondol→ 'gondola'	119.7
* γondol- 'gondola'	119.0
MAX(R)	118.7
MAX(A)	111.8
* aγor- 'market'	117.2
* áγor- 'market'	116.9
DEP(A)	109.4
* -on 'Gen.Pl'	108.9
* -á 'Nom.Sg'	107.606
* ←on 'Gen.Pl'	104.3
* ←a 'Nom.Sg'	102.1
DEP(R)	98.5
* -a 'Nom.Sg'	90.3
* -ón 'Gen.Pl'	86.9
* aγor→ 'market'	65.8
* γóndol- 'gondola'	61.1

For the meaning 'gondola-Gen.Pl', all candidates with a post-stressing or an unstressed root are ruled out by the high-ranked lexical constraints that militate against these forms. All candidates with stress on the affix in the surface form are ruled out by the high-ranked MAX constraints. The selected surface form is always /γóndolon/, therefore I exclude for brevity's sake the candidates with the alternative surface forms, because they never win anyway (as a result, MAX(R) and DEP(A) are never violated in the tableau). The candidates with an underlyingly stressed suffix are ruled out not because of the lexical constraint *|-ón| 'Gen.Pl' that militates against it, but because of MAX(A), which outranks the lexical constraint militating against an underlyingly *unstressed* suffix, *|-on| 'Gen.Pl'.

⁶⁷ For space reasons, I only list the constraints I discuss in tableaux (208) and (211).

(208) Meaning-to-form evaluation

'gondola-Gen.Pl'	* γondol→ 'gondola'	* γondol- 'gondola'	MAX(R)	MAX(A)	DEP(A)	* -on 'Gen.Pl'	* ←on 'Gen.Pl'	DEP(R)	* -ón 'Gen.Pl'	* γóndol- 'gondola'
☞ 84.9%						*				*
γóndol+on /γóndolon/						*		*		
γondol+on /γóndolon/		*!				*		*		
γondol→+on /γóndolon/	*!					*		*		
15.1%				*!					*	*
γóndol+ón /γóndolon/				*!				*	*	
γondol+ón /γóndolon/		*!		*				*	*	
γondol→+ón /γóndolon/	*!			*				*	*	
γóndol+←on /γóndolon/				*!		*		*		*
γondol+←on /γóndolon/		*!		*		*		*		
γondol→+←on /γóndolon/	*!			*		*		*		

Due to the probabilistic ranking of constraints in the GLA, the ranking between MAX(A) and *|-on| 'Gen.Pl' can switch from time to time. Therefore, the winning candidate under the ranking in tableau (208) is in 84.9% of the cases |γóndol+on| /γóndolon/, with an underlyingly unstressed affix, but in 15.1% of the cases |γóndol+ón| /γóndolon/, with an underlyingly stressed affix. This happens to be the form that I (and other linguists) have assumed for Modern Greek in table (177) and (178). In the evaluations where the underlying form |γóndol+ón| is chosen, *|-on| 'Gen.Pl' outranks MAX(A). Looking at the ranking of lexical constraints separately, we can see that the constraints clearly decide for one underlying form, |γóndol+ón|:

(209) Separate ranking of lexical constraints

{*|γondol→| 'gondola', *|γondol-| 'gondola'} >> *|γóndol-| 'gondola'
 {*|-on| 'Gen.Pl', *|←on| 'Gen.Pl'} >> *|-ón| 'Gen.Pl'

Looking at the ranking of faithfulness constraints separately, I can state that the ranking I stipulated in (186) is also achieved:

(210) Separate ranking of faithfulness constraints

$$\text{MAX(R)} \gg \text{MAX(A)} \gg \text{DEP}$$

A case of a post-stressing form is provided in tableau (211). Due to the ranking of $\text{DEP(A)} \gg *|-\acute{a}|$ 'Nom.Sg' in 73.9% of the cases, the candidate with an underlyingly post-stressing root and an underlyingly stressed suffix $|a\gamma or \rightarrow \acute{a}| / a\gamma or \acute{a} /$ wins. In 26.1% of the cases, the constraints swap places in the hierarchy and the candidate with the underlyingly post-stressing root and the underlyingly unstressed suffix $|a\gamma or \rightarrow -a| / a\gamma or \acute{a} /$ wins.

(211) Virtual production of *aγora*

'market-Nom.Sg'	MAX(R)	* aγor- 'market'	* áγor- 'market'	MAX(A)	DEP(A)	* -\acute{a} 'Nom.Sg'	* ←-a 'Nom.Sg'	DEP(R)	* -\acute{a} 'Nom.Sg'	* aγor→ 'market'
aγor-a /aγorá/		*!			*				*	
áγor-a /aγorá/	*!		*		*				*	
26.1 % aγor→-a /aγorá/					*!				*	*
aγor-\acute{a} /aγorá/		*!				*				
áγor-\acute{a} /aγorá/	*!		*			*				
☞ 73.9% aγor→-\acute{a} /aγorá/						*				*
aγor←-a /aγorá/		*!		*	*		*			
áγor←-a /aγorá/	*!		*	*	*		*			
aγor→-←-a /aγorá/				*!	*		*			*

Again, if we separate the lexical constraints from the faithfulness constraints, we get the underlying forms as assumed in table (178).

(212) Separate ranking of lexical constraints for *aγora*

$$*|a\gamma or-| \text{ 'market' } \gg *|\acute{a}\gamma or-| \text{ 'market' } \gg *|a\gamma or \rightarrow| \text{ 'market' }$$

$$*|-\acute{a}| \text{ 'Nom.Sg' } \gg *| \leftarrow -a| \text{ 'Nom.Sg' } \gg *| -a| \text{ 'Nom.Sg' }$$

This result is only partially satisfying; the learners end up with allomorphy: for a given meaning, they choose for one underlying form, e.g. $|\gamma \acute{o}ndol+on|$ in one evaluation, but for $|\gamma \acute{o}ndol+\acute{o}n|$ another time. This does not affect their communication skills, since the surface forms are uniform, but violates a

principle such as lexical economy. A possible solution to the problem could be to split up the virtual production step into the computation of the underlying form from meaning as a first step, and subsequently the computation of surface form from the computed underlying form as a second step (Paul Boersma, p.c.). Another solution could be to split up the acquisition process: first, the mapping from surface form to underlying form could be modelled, and then the mapping from underlying form to meaning.

The next step will push the model one step forward: section 6.6 shows how both underlying forms and surface forms can be learned, given meaning and overt forms.

6.6 Modelling comprehension: two levels of hidden structure

This section shows how cases with two hidden levels of representations can be modelled. I now combine the two approaches of learning surface forms, as addressed in step 2 of the introduction, and learning underlying forms, as addressed in step 3 of the introduction, resulting in step 4 of the introduction. Imagine that a mother is pointing out a boat to her child, pronouncing the overt form [γόνδολα] at the same time: thus I stipulate that meaning and overt form are given to the learner, and that both a surface form (in the form of /(\gammaón.do)la/) and an underlying form (in the form of |\gammaóndol+a|) have to be created, as was illustrated in figure (173).

6.6.1 The training data

The training data are given in (213). I will limit myself to the two-way contrast in morphemes, in order to avoid an explosion of candidates and constraints. From the stress in the overt form, the learner has to infer foot structure in the surface forms and stress marks in the underlying forms.

(213) Training data

[γόνδολα]	[θάλασα]	[astráγalos]
[γόνδολον]	[θalasón]	[astrayalón]

The learner does not know whether the language at hand has lexical or grammatical stress, therefore three kinds of constraints are involved: faithfulness constraints, structural constraints, and lexical constraints. The faithfulness constraints are as in (196), the lexical constraints are as in (197), and the structural constraints are as in (179). For the moment, all representations will be processed in a parallel fashion; the candidates for evaluation are therefore quadruplets consisting of meaning / underlying form / surface form / overt form.

6.6.2 The pool of underlying forms to choose from: GEN

The possible underlying morphemes that GEN provides are listed in (214). For the purposes of this section I only assume a two-way contrast in the underlying forms.

(214) Possible underlying morphemes

γondol-	θalas-	astrayal-	-a	-on	-os
γóndol-	θálas-	astráyal-	-á	-ón	-ós

GEN also has to provide possible surface forms, this time candidates with varying foot structures:⁶⁸

(215) Possible surface structures

/γόν.do)la/	/(θά.la)sa/	/(á.stra)γα.los/	/(á.stra)γα.lon/
/γόν.do.la/	/(θά)la.sa/	/(á)stra.γα.los/	/(á)stra.γα.lon/
/γον(do.lá)/	/θα(la.sá)/	/a(strá.γα)los/	/a(strá.γα)lon/
/γον.do(lá)/	/θα.la(sá)/	/a(strá)γα.los/	/a(strá)γα.lon/
/(γόν.do)lon/	/(θά.la)son/	/a.stra(γα.lós)/	/a.stra(γα.lón)/
/(γόν)do.lon/	/(θά)la.son/	/a.stra.γα(lós)/	/a.stra.γα(lón)/
/γον(do.lón)/	/θα(la.són)/		
/γον.do(lón)	/θα.la(són)/		

⁶⁸ GEN is pretty restricted here: there are no secondary stresses, i.e. more than one foot, in the candidates, and stress can never occur on the second syllable of the root.

6.6.3 The constraint set

The constraints are once more the faithfulness constraints of (196), plus the structural constraints of (179), plus lexical constraints militating against each possible underlying morpheme of (197) with additional constraints against ‘ankle’-forms:

- (216) Additional lexical constraints:
- *|astrayal-| ‘ankle’
 - *|astráyal-| ‘ankle’
 - *|-os| ‘Nom.Sg.M’
 - *|-ós| ‘Nom.Sg.M’

6.6.4 Results: the chosen underlying forms

The virtual learners learned as shown in the tableaux in sections 3.3 to 3.5, applying the GLA reranking strategy. They are successful with respect to the overt forms they learn to produce: these are the same than they were fed in the training phase. Communication is therefore guaranteed. However, surface forms and underlying forms are not always unique, as we can see from table (217). The meaning ‘gondola-Nom.Sg’ takes always the underlying form |γόνdol+a|, the surface form /(\γόν.do)la/, and the overt form [γόνdola], throughout all learners. For all the other forms, there is variation, but only in one of the two hidden forms, either in the surface form or in the underlying form. For instance, the meaning ‘gondola-Gen.Pl’ always takes the surface form /(\γόν.do)lon/, but varies between the underlying forms |γόνdol+on| and |γόνdol+ón|. The meaning ‘sea-Gen.Pl’ always takes the underlying form |\thalas+ón|, but varies between the surface forms /\thetaa.la(són)/ and /\thetaa(la.són). We can observe that for the root |γόνdol-|, the stressed underlying form is uniformly chosen. For the suffix |-a|, always the unstressed form is chosen; the same with the suffix |-os|. This is because there is no alternation in the overt forms: The root |γόνdol-| is always stressed, and the two suffixes are always unstressed. The morphemes that have alternating stress in the overt forms display allomorphy.

However, it is not the case that overt forms are always faithfully mapped onto underlying forms. If underlying forms were completely faithful to the corresponding surface forms, it would imply that the genitive plural suffix *-on*, when combined with the root | $\gamma\acute{o}ndol$ -|, would always take the underlyingly unstressed form |*-on*|, but when combined with the root | $\theta\acute{a}las$ -|, would always take the underlyingly stressed form |*-ón*|. This is not the case; in many (not necessarily the majority) of the cases, the ‘correct’ underlying form (i.e. the underlying form I have assumed for Modern Greek in section 6.2) is chosen.

The problem is once again the interference of faithfulness constraints with lexical constraints, as we can see in (220). $MAX(A)$ outranks $*|{-on}|$ ‘Gen.PI’, and by that rules out the candidate that we would prefer. If the two constraints switch their place in the hierarchy at evaluation time (as they do from time to time, because they are ranked close to each other), the candidate I prefer would be chosen: | $\gamma\acute{o}ndol+ón$ |. If we separate the ranking of the lexical constraints from the faithfulness constraints, we can see that the selection of underlying forms is uniform:

- (218) Ranking of lexical constraints
 $*|\gamma\acute{o}ndol$ -| ‘gondola’ \gg $*|\gamma\acute{o}ndol$ -| ‘gondola’
 $*|{-on}|$ ‘Gen.PI’ \gg $*|{-ón}|$ ‘Gen.PI’

- (219) Ranking of faithfulness constraints
 $MAX(R)$ \gg $MAX(A)$ \gg DEP

Tableau (220) shows the evaluation of $\gamma\acute{o}ndolon$ with the relevant constraints (i.e. the lexical constraints on other morphemes than $\gamma\acute{o}ndol$ - and *-on* are excluded).

I furthermore tested whether different initial rankings such as $*LEX \gg \{ FAITH, STRUCTURAL CONSTRAINTS \}$ or $\{ *LEX, STRUCTURAL CONSTRAINTS \} \gg FAITH$ would lead to uniform underlying forms and uniform surface forms, but they did not improve learning. During the learning process, faithfulness is needed. I also tested whether other input frequencies would change the result, but they did not. As already addressed in section 6.5, it will be worth investigating whether a more serial approach (modelling first perception, then recognition; or modelling the mapping from meaning to UF and then the mapping from UF to SF in production) instead of the fully parallel approach would make a difference.

In the following sections I discuss some alternatives to the on-line learning approach of underlying forms I proposed here.

6.7 Alternative approaches to the learning of underlying forms

In section 6.7.1 I discuss how Constraint Demotion (Tesar 1995) fares in my on-line learning approach, as opposed to the GLA. Section 6.7.2 discusses Lexicon Optimization as a means to determine underlying forms. Section 6.7.3 explores the off-line learning approach *inconsistency detection and surgery* by Tesar et al. (2003). Section 6.7.4 outlines probabilistic unsupervised learning of underlying forms (Jarosz 2006) that makes use of the Expectation Maximization Algorithm (Dempster et al. 1977).

6.7.1 Learning underlying forms with Constraint Demotion

To compare whether CD (Tesar 1995) would fare better as a reranking strategy in than the GLA, I ran a simulation with 10 virtual learners for the two-way contrast outlined in section 6.4 where everything was kept the same except for the reranking strategy. It turned out that the 10 CD learners arrived at a ranking that rendered the correct SFs in the production step. However, they decided to create faithful lexical allomorphs: instead of choosing just one morpheme for each meaning, as the GLA learners in their place did, they sometimes chose two, namely in the cases where the surface forms yielded alternation. The root |θálas-| occurred as consistently underlyingly unstressed when combined with the affix |-ón|, and as consistently underlyingly stressed when combined with the affix |-a|. The affix |-on| occurred as underlyingly stressed when combined with the root |θálas-|, and as underlyingly unstressed when combined with |γóndol-|:

(221) The resulting lexicon of CD learners:

γóndol-	θálas-	-a	-on
	θálas-		-ón

This was due to the fact that the CD learners were able to establish a ranking between the lexical constraints, but failed to rank the faithfulness constraints:

- (222) The ranking of an example CD learner:
 {MAX(R), MAX(A), DEP(R), DEP(A), *|-á| ‘Nom.Sg’, *|ɣondol-| ‘gondola’}
 >>
 { *|-a| ‘Nom.Sg’, *|-on| ‘Gen.PI’, *|θálas-| ‘sea’, *|ɣóndol-| ‘gondola’ }
 >>
 { *|θálas-| ‘sea’, *|-ón| ‘Gen.PI’ }

This is a possible solution; however, it is not the most restrictive lexicon that can be found (if you assume that there can only be one underlying form for every lexical item, as listed in (178) of section 6.2). It means that in the case of alternation, SFs are always faithfully mapped onto UFs. If every encountered item is stored as is in the lexicon, large parts of the grammar can be considered superfluous. Like in the case with the GLA, it is worthwhile investigating whether a more serial processing approach can make a difference.

6.7.2 Lexicon Optimization

In Lexicon Optimization, as proposed in Prince & Smolensky (1993:191) and further developed in Itô et al. (1995), the underlying form of a word is determined by evaluating different possible underlying forms with respect to a surface form which is optimal in the ranking of the language. The optimal surface form is determined by the ranking of structural constraints, and the appropriate underlying form for this surface form is determined by faithfulness: the most faithful underlying-surface pair is the most harmonic one, and chosen as the optimal pair. Applying this to the case in point, Modern Greek, it shows that Lexicon Optimization is problematic: even given the language’s constraint ranking, the decision between two underlying forms, the stressed |ɣóndol+on| and the unstressed |ɣondol+on|, cannot be made by the grammar, as shown in tableau (223).

(223) Lexicon Optimization in Modern Greek

UF	SF	MAX(R)	MAX(A)	NONFIN	AFR	AFL	FTBIN
☞ γόνδολ+ον	/(\γόνδο)lon/				*		
γόνδολ+ον	/\γονδο(λόν)/	*!		*		**	*
☞ γονδολ+ον	/(\γόνδο)lon/				*		
γονδολ+ον	/\γονδο(λόν)/			*!		**	*
γόνδολ+όν	/(\γόνδο)lon/		*!		*		
γόνδολ+όν	/\γονδο(λόν)/	*!		*		**	*
γονδολ+όν	/(\γόνδο)lon/		*!		*		
γονδολ+όν	/\γονδο(λόν)/			*!		**	*

The difference could be made by DEP(R), which would decide that the unstressed underlying form is the optimal one. This would put the root *γονδολ-* on a par with the root *θάλας-*, which is underlyingly unstressed, too. Stress in *γόνδολον* would then be the result of the phonological default assignment. However, when combining *γονδολ-* and *θάλας-* with *-ον*, the difference between the two roots becomes apparent: *γονδολ-* maintains stress (becoming *γόνδολον*), while *θάλας-* loses it to *-ον* (becoming *θάλασόν*). The lexicon-optimization process is therefore not sufficient.

Tesar & Smolensky (1996, 2000:77) extend Lexicon Optimization by combining the evaluation of different input-output pairs for a given word with input-output pairs of different words, i.e. including paradigmatic comparison. Paradigmatic comparisons involve off-line learning: the learner gathers informative data until she changes her grammar. I argue that this is problematic in a more realistic learning situation, because the learner does not know when she has gathered enough data, and whether these data are informative. This is not the case for an on-line learning algorithm as the GLA: a GLA learner will adjust her grammar as long as incoming forms trigger an error detection. If incoming forms do not trigger any error detection any more, learning stops, and the grammar is not changed any further.⁶⁹

⁶⁹ The learning curve of the constraints reaches a plateau in this case, and the constraints do not alter their ranking values any more.

6.7.3 Comparison to inconsistency detection and surgery

Inconsistency detection and surgery (Tesar et al. 2003) makes use of paradigmatic comparison to determine underlying forms. The point of departure in Tesar et al.'s approach (2003) was similar to the one in section 6.4: surface forms are given to the learner, as is the morphemic distinction between roots and affixes. As a first step, the learner tries to find a ranking for the language data. If there is no such ranking, meaning that there is inconsistency in the data, she will modify her lexicon. To be able to do so she will gather paradigmatic information. Forms that do not show any alternation in the paradigm are faithfully mapped onto underlying forms. All alternating forms are listed as underlyingly unstressed. She will then randomly modify one of the unstressed forms and try again to find a ranking for the data. If she does not find one, she will reset the modified form to unstressed and try to modify another form she listed as unstressed. She will proceed with modifying lexicon and ranking in turns until she finds a lexicon consistent with a ranking. Via *surgery* the learner is able to remember which pairs of forms have already been tested and discarded. This algorithm proceeds in an off-line fashion: when no ranking is found that is consistent with all the data, the lexicon is modified, after gathering all possible surface forms. But a learner does not know when she gathered enough data to go on with learning. Furthermore, she has to remember all the forms she discarded as not optimal.

6.7.4 Considering multiple grammars

Another approach to the learning of underlying forms is considering multiple grammars (in fact, *all* grammars possible) at a time, as in Jarosz (2006). It makes use of the Expectation Maximization Algorithm (Dempsey et al. 1977) applied to Optimality Theoretic grammars. The proposed approach is similar to the one proposed here in 6.4: all possible underlying forms and all rankings of constraints are initially equally probable. The probability of the underlying forms and the probability of the rankings are computed in combination with the probability of observed forms.

In this approach, the probability of all possible constraint rankings and underlying forms is computed, consulting the distribution of surface forms iteratively. Every iteration step consults all possible grammars. This means

that, at all times in learning, all constraint rankings (i.e., all possible grammars $N!$) are present. This does not scale well: a constraint set of e.g. 30 constraints yields $30!$ possible grammars: every constraint added to the possible set of UG lets the number of possible grammars grow exponentially. While this might be a proper mathematical model for finding rankings and underlying forms given a distribution of surface forms, it is not suitable as a learnability approach.

6.7.5 A note on Richness of the Base

What I proposed here is something which some OT-ists think is forbidden: I put restrictions on what is generally referred to as the *input* in OT, namely the underlying form. According to Prince & Smolensky's (1993) principle *Richness of the Base*, there should be no restrictions on the input, but at the same time they acknowledge the need for concrete underlying forms. As Tesar & Smolensky (2000:30) put it:

- (224) Tesar & Smolensky's definition of Richness of the base
"Richness of the base: The set of possible inputs to the grammars of all languages is the same. The grammatical inventories of languages are defined as the forms appearing in the outputs that emerge from the grammar when it is fed the universal set of all possible inputs (P&S section 9.3)."

This means that just as constraints are universal in OT, so is the pool of possible underlying forms. This does not mean that a given language makes use of all possible underlying forms, just as little as it makes use of all constraints. The pool of underlying forms I assumed for Modern Greek in e.g. (195) is only a small part of a universal set of possible underlying forms. A child learning Chinese could choose from the forms in (195), but probably wouldn't pick any of them, because they do not relate to any form she is exposed to in her language. So what 'input' in production-directed OT refers to is the set of possible underlying forms. For the lexicon of a given language it is necessary to make a selection among these possible forms. Prince & Smolensky (1993) proposed Lexicon Optimization as the mechanism for selecting concrete underlying forms: pairs of possible

underlying forms and surface forms (since the grammar evaluates just one surface form as correct, the surface form always stays the same in the pairs, and only the underlying forms vary) are compared to each other. The most harmonic pair that is chosen by the constraint ranking contains the best underlying form for that surface form. As I argued in section 6.7.2, Lexicon Optimization can lead to abundant allomorphy: the underlying forms for given surface forms are the most faithful ones. This is undesirable, because if the underlying forms are always the most faithful ones, we end up with having an underlying form for each different surface form (e.g. German [tak] for ‘day-Sg’, and [tag+ə] for ‘day-Pl’). This entails a lot of redundant information in the lexicon, which one would like to exclude. Itô et al. (1995) and Tesar & Smolensky (2000) got around that problem by including paradigmatic comparison. I already argued that this implies off-line learning, which I want to discard as a possibility. Instead of assuming that the pool of possible underlying forms is made available by Richness of the Base, I argue that it is Freedom of Analysis, and therefore GEN, which provides the pool. Underlying forms are an *output*, and the choice for an underlying form over another is then made by constraint evaluation. I still embrace the idea of Richness of the Base. In my account, meaning and overt forms are inputs, and are therefore the *rich base*: overt forms are the input to the comprehension process (Boersma 2000), and meaning is the input to the production process.

6.8 Discussion

The alternatives discussed in 6.7.1 to 6.7.4 have in common that they are all off-line learning approaches: the learner first has to gather paradigmatic information before she can begin to modify her lexicon. This is not a very natural approach to language acquisition. It is not clear at what point the learner knows that she gathered enough information and will not encounter any further alternations. Moreover, she has to maintain access to all the observable forms or even to all possible grammars at all times in learning. This implies unlikely mnemonic processes: incoming forms are not processed and then discarded, but stored for later consultation. Furthermore, the concept of surgery relies on backtracking, which might not be possible in a real learning situation. The multiple-grammar approach of Jarosz (2006)

relies on the availability of all possible grammars during the whole learning process, which is not very realistic to assume, either.

In this chapter, it has been shown how underlying forms can be learned by a rather “stupid” on-line learning algorithm that takes meaning into account. An on-line approach of learning is better than an off-line approach, because one form is processed at a time, under one ranking at a time. Former processed forms or rankings do not have to be remembered, because their occurrence is implicitly stored in the ranking of the constraints. The ranking is adjusted systematically. No extra learning mechanisms are required than the ones already involved in a general grammar learning model of OT (e.g. Boersma 1997): interpretation of incoming forms and constraint reranking as a reaction to error detection. The learning of underlying forms takes place by learning the grammar. This resolves the problem of whether it is the grammar or the lexicon that has to be learned first.

Why is it feasible to only have a few underlying forms and not many? If underlying forms would always be faithful to surface forms, this would render the concept of a grammar superfluous: everything would be in the lexicon. Psycholinguistic evidence furthermore indicates that words are composed of and stored as parts (e.g. roots and suffixes) in the lexicon (e.g. Chomsky & Halle 1968:12), and not as forms as a whole. This means that speakers of a language should be able to decompose incoming forms and find connections between them. In the proposed model, this is ensured by linking forms to meaning in the learning process. The learning approach of underlying forms makes use of grammatical restrictions on the lexicon in form of lexical constraints. This means in effect that a strict demarcation between grammar and lexicon cannot be upheld; the lexicon becomes a part of the grammar.