Acta Linguistica Hungarica, Vol. 57 (4), pp. 458–473 (2010)
DOI: 10.1556/ALing.57.2010.4.7

DOES TURKISH DISS HARMONY?*

MARKUS ALEXANDER PÖCHTRAGER

Fen-Edebiyat Fakültesi
Boğaziçi Üniversitesi
Batı Dilleri ve Edebiyatları Bölümü
34342 Bebek/İstanbul
Turkey
markus.pochtrager@boun.edu.tr

Abstract: This article presents a Government Phonology (GP) analysis of disharmonic words in Turkish. According to GP, phonology is exceptionless. Following this claim, I will argue that the generalisations intended to capture vowel harmony in Turkish had been stated in the wrong way, leading to disharmonic words as an artefact of a faulty analysis. Once this is remedied, the exceptions vanish, allowing for a unified treatment of harmonic and disharmonic words. This also takes into account further details of the Turkish vowel system which had not been incorporated in previous analyses, as well as distributional asymmetries between stems and suffixes.

Keywords: Turkish, vowel harmony, exceptions, elements, Government Phonology

1. Introduction

1.1. Exceptions and GP

"The exception proves the rule," they say. There is a deeper truth in that seemingly non-sensical statement, namely that *exception* is a rela-

* Thanks to hip hop music for creating the slang clipping *to dis(s)* from *to disrespect*. This article is somewhat anachronistic—at the time of writing, the theory of Government Phonology is undergoing a major overhaul. The improved version currently being developed, GP 2.0 (Pöchtrager 2006; 2009; Živanović–Pöchtrager 2010), still presents us with some very fundamental questions, in particular in the area of element theory, which is why this article (mostly) keeps to the old model.

tive notion: If there is no rule, there can be no exception to it; and the nature of the exception depends on the nature of the rule. In this article I will look at Turkish vowel harmony (TVH) and exceptions to it, i.e., disharmonic words; words that violate harmony. My analysis is couched in the framework of Government Phonology, GP (Kaye et al. 1985; 1990; Kaye 1990), which takes a very clear stand on exceptions—it is built on the assumption that there aren't any, that phonology is exceptionless. Proceeding from that assumption, I will claim that those so-called exceptions to TVH are really an illusion, that we only think of them as such because the generalisations intended to capture TVH have been stated in the wrong way. Once this is corrected, the exceptions vanish. The modification of our view on TVH will also help us to understand another property of the Turkish vowel system better, namely that there are more vowels (nuclear expressions) than assumed so far. We begin by considering one assumption that lies at the very heart of GP: the Minimality Hypothesis (MH), following Kaye (1992, 141).

(1) **Minimality Hypothesis (MH):**
Processes apply whenever their conditions are met.

The MH is very clear about exceptions: There should be none. If a phonological process **can** take place, it **must** take place. If in a certain language **I** spreads, then **I** should spread every time it occurs. There is no way that certain words are marked as exceptions to spreading, as this would be a violation of the MH. With this in mind, let us now look at how the regular (non-exceptional) cases of TVH have been analysed in the GP literature. This will furnish the necessary background on the basis of which we can then assess the so-called exceptions to TVH.

1.2. The well-behaved part of TVH

TVH is hardly an understudied phenomenon. (See references in quoted literature.) Also within GP there have been several proposals of how TVH should be analysed. The chart in (2) gives the harmonic combinations of (nuclear) phonological expressions (PEs) in TVH, following Charette–Göksel (1994; 1996); Denwood (2002); Kaye (2001).[1]

[1] Those texts do not all agree on which element is the head (the underlined element). I will follow Charette–Göksel (1996) here.

In the present section we will have a look at the key proposals of the analysis given in Charette–Göksel (1996).

(2)

Stems	PEs		Ablative Lexically $(\{\}\underline{\mathbf{A}})$	2.sg Possessive Lexically empty
kıl	$(\{\}\_\_)$	'hair'	*kıl-dan*	*kıl-ın*
dal	$(\{\}\underline{\mathbf{A}})$	'branch'	*dal-dan*	*dal-ın*
kul	$(\{\}\underline{\mathbf{U}})$	'subject'	*kul-dan*	*kul-un*
kol	$(\{\mathbf{A}\}\underline{\mathbf{U}})$	'arm'	*kol-dan*	*kol-un*
kil	$(\{\}\underline{\mathbf{I}})$	'clay'	*kil-den*	*kil-in*
kel	$(\{\mathbf{A}\}\underline{\mathbf{I}})$	'bald patch'	*kel-den*	*kel-in*
kül	$(\{\mathbf{I}\}\underline{\mathbf{U}})$	'ash'	*kül-den*	*kül-ün*
göl	$(\{\mathbf{I},\mathbf{A}\}\underline{\mathbf{U}})$	'lake'	*göl-den*	*göl-ün*

Turkish has eight different PEs that can occur in the initial nucleus of Turkish words.[2] These eight different PEs are created by licensing constraints (LCs) as discussed in Charette–Göksel (1996). All expressions but one are headed: *ı* [ɨ] is the realisation of a (non-p-licensed) empty nucleus, i.e., its melodic representation is $(\{\}\_\_)$, with both head and operator position empty. All other PEs have at least the head position filled. In non-initial position, e.g., in suffixes, only two of the eight nuclear PEs occur: $(\{\}\underline{\mathbf{A}})$ or $(\{\}\_\_)$. An example of a suffix with $(\{\}\underline{\mathbf{A}})$ is the ablative suffix *-dan*. This suffix alternates with *-den*. The PE $(\{\}\_\_)$ occurs in the possessive suffix of the second person singular, *ın*, which alternates with *in*/*un*/*ün*.

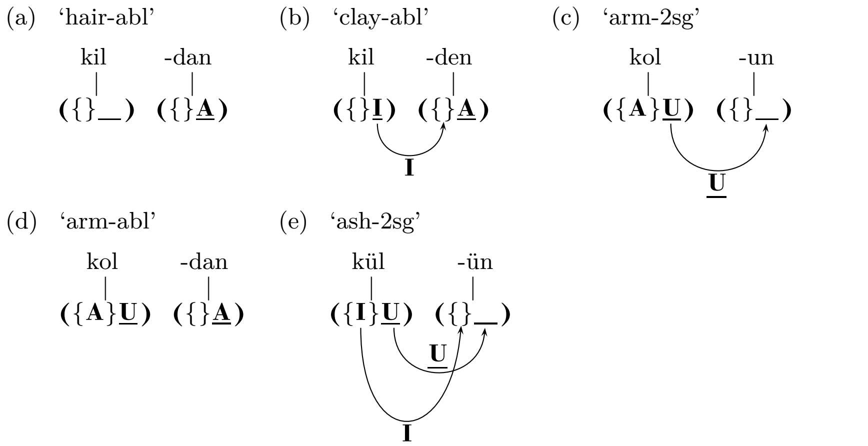
TVH is usually understood as **I** and/or **U** spreading to the right. This leads to an alternation in the suffixes as seen in (2). Some examples from (2) are given in more detail in (3). In (3a), the stem contains neither **I** nor **U**, hence nothing spreads and the suffix stays unchanged. The **I** in the stem of (3b) spreads to the operator position of the suffix, changing *-dan* to *-den*.[3] In (3c) **U** spreads into the head position of the underlying empty nucleus and we derive *-un*. (3d) is slightly different in that **U** fails to spread if it has to give up its role as a head, i.e., when the target

[2] This number will be subject to further qualification, as we shall see.

[3] Note that the *e* of the suffix, $(\{\mathbf{I}\}\underline{\mathbf{A}})$, differs from the *e* in the stem in (2), $(\{\mathbf{A}\}\underline{\mathbf{I}})$. Charette and Göksel (1996) justify this by pointing out that Turkish has two different *e*'s. While they are right in that Turkish has more than one *e*, they are wrong with respect to where the different *e*'s occur. I will not go into this problem of their analysis as I am ultimately going to reject it anyway.

position is already headed.[4] (3e) illustrates the spreading of **I** and **U** at the same time. The target position is empty, so **U** can spread to the head position and **I** to the operator. The result is *-ün*.

(3)



TVH applies within the word, but not across a compound boundary. In a compound like *demirkapı* [[*demir*][*kapı*]] 'iron door' the **I** in the *e* of *demir* spreads to the right within the word, giving *i* (*demi̲r*), but cannot move on to the second noun (**demirkepi*). Each member of the compound forms a domain of its own that is inaccessible from outside.[5] As we shall see shortly, this property of compounds was exploited in the past to account for exceptions to TVH.

2. The problem

The problem for the analysis so far comes in the shape of exceptions to TVH, also known as "disharmonic words" (DWs). There are two types of exceptions. The first type of exception has to do with a failure to spread. In section **1** we saw that **I** and **U** spread to the right, with **U**, but not **I**, being subject to certain phonological restrictions. According to the MH

[4] For Charette and Göksel (1996) this follows from the LCs generating the vowel system.

[5] Following Kaye (1995), who introduced the two functions concat( ) ("concatenate") and φ( ) ("do phonology"), we have φ(concat(φ(*demir*),φ(*kapı*))) here. The φ( )-function seals off its argument for **I**/**U**-spreading from outside: *kapı* is in its own domain (since it is the argument of the φ( )-function) and hence inaccessible.

given in (1), this should be true of any word containing **I** or **U**. This is clearly violated in words such as *bira* ‘beer’, *kitap* ‘book’, *elma* ‘apple’, *dünya* ‘world’, *mühim* ‘important’, *kilo* ‘kilo’ etc. In *bira* ‘beer’ we have the element **I** in *i* **(**{}**I̲)** and this **I** should spread to the right, but fails to do so, as indicated by a broken arrow and a question mark in (4a).

(4) (a) bira ‘beer’ (b) hafif ‘light’

The second kind of exception is illustrated by words such as *hafif* ‘light’, *arzu* ‘wish’, *kapsül* ‘capsule’, *kısmi* ‘partially’ etc. Charette and Göksel (1996) propose that non-initial positions can only contain **(**{}**A̲)** or **(**{}**__)**, which predicts that any **I** or **U** we find in a non-initial position must have come from the left. In *hafif* ‘light’ (4b) the **I** in **(**{}**I̲)** should have come from the left, but it cannot have as the initial nucleus does not contain **I**.

Both types of exception also occur together: In *cesur* ‘brave’ the **I** in the first nucleus should spread, but it does not; the **U** in the second nucleus should come from the left, but it does not.

The second type of exception is easier to get a handle on. Words like *hafif* ‘light’ violate a language-specific assumption about non-initial positions which might as well be wrong. It could well be the case that suffixes are restricted in what kind of nuclear PEs they can contain, but that non-initial nuclei in stems are not subject to the same restrictions.[6]

The first type of exception is more problematic, however. A word like *bira* ‘beer’ violates a universal principle, viz., the MH. The MH is one of the core parts of GP, and if we had to acknowledge that it makes wrong predictions, we would be forced to give up a major part of the theory. Clearly we will have to find a way out of this predicament.

3. Attempts at a solution

In order to come to terms with DWs, Denwood (1998) and Kaye (p.c.) have claimed that such DWs involve dummy morphology, i.e., that they

[6] In fact, this is exactly the proposal I will make in section **6**.

have the morphological structure of compounds even though the individual parts have no meaning by themselves.[7] A DW such as *bira* 'beer' is assumed to consist of two parts, *bi* and *ra*, which together make up [[*bi*][*ra*]]. Given that TVH always stops at compound boundaries, the lack of harmony is to be expected.

Such an analysis has an immediate advantage beyond rescuing us from a violation of the MH and explaining the existence of DWs: The compound structure also explains the irregular stress in *bíra* 'beer', whose (initial) stress pattern is like that of compounds, cf. a "regular" ("non-dummy") compound such as [[*demír*][*kapı*]]. Compound stress is practically always on the first member,[8] and if *bira* is really [[*bi*][*ra*]], then we predict the first nucleus to bear stress, which is correct. The dummy morphology approach seems to kill two birds with one stone.

There is, however, a fly in the ointment, or rather: three flies. The dummy morphology analysis makes three (problematic) predictions; problematic in so far as they are not borne out by the facts. Let us consider these predictions in turn:

Firstly, the individual parts (*bi* and *ra*) of such dummy compounds should always be big enough to be domains in their own right (P1). Minimal words in Turkish have the shape (C)VC (*can* 'soul', *top* 'ball', *ev* 'house') or (C)Vː (*dağ* [daː] 'mountain'). (C)V by itself is too small and usually restricted to closed word classes (*bu* 'this', *ne* 'what', *ki* 'that', *o* 'he/she/it'), with only three exceptions (such as *su* 'water'). In other words, if *bi* and *ra* were the two domains that together made up *bira*, they would be very unusual, if not ungrammatical as domains.

Secondly, if DWs are compounds, they should all have compound stress, i.e., on the first member (P2). Again, this is not what we find: *elmá* 'apple', *meyvá* 'fruit', *dünyá* 'world', *sahíl* 'beach', *kitáp* 'book' etc. are all DWs but have final stress. If they were dummy compounds,

[7] For morphology in GP, cf. Kaye (1995). Denwood (1998) considers Turkish words to have a rather complex internal structure in general. For her all words longer than (C)VC or (C)VV consist of a sequence of several templates (domains) which are combined analytically, with a given template being dependent on the previous one, i.e., [[A] B], [[[A] B] C] etc. For DWs Denwood employs **in**dependent analytic morphology of the type [[A][B]], which is an addition to handle DWs and does not fall out from her model of templates.

[8] The second member can only ever bear stress when it is a verb, e.g., in *bilgisayár* 'computer' from *bilgi* 'knowledge' and *sayár* '(something) that counts', cf. Göksel – Kerslake (2005, 28–9). This, of course, is hardly applicable to the dummy morphology claimed to exist in DWs.

we should expect ***élma*, **méyva* etc. Furthermore, as in all Turkish words that bear final stress, this final stress will move further to the right if suffixes are attached: *kitáp* 'book', *kitap-lár* 'books', *kitap-lar-ím* 'my books', *kitap-lar-ım-dá* 'in my books', etc. Again, this is in marked contrast to (real) compounds: The stress in *demírkapı* 'iron door' will never move: *demírkapı-lar* 'iron doors', *demírkapı-lar-ım* 'my iron doors', *demírkapı-lar-ım-da* 'in my iron doors', etc.

Thirdly, DWs should allow for **any** vowel combination (P3). There are no restrictions on what kind of vowel combinations can occur between the two members of a compound; hence we should also expect not to find any such restrictions in DWs. And again, this is not borne out by the facts. DWs do not allow for any vowel combination. Crucially, empty nuclei are always harmonised.[9] The realisation of an empty nucleus, *ı* [ɨ], never follows any other vowel but *a* or *ı* (which is a harmonic combination, cf. the chart in (2)). Consider the following chart of combinations in words with two realised nuclei. (This is extendable to longer sequences.)[10]

(5)

		N_2							
		ı	*a*	*u*	*o*	*i*	*e*	*ü*	*ö*
N_1	*ı*	h	h	∗	✓	✓	✓	∗	∗
	a	h	h	✓	✓	✓	✓	✓	✓
	u	∗	h	h	✓	✓	✓	✓	✓
	o	∗	h	h	✓	✓	✓	∗	✓
	i	∗	✓	✓	✓	h	h	✓	✓
	e	∗	✓	✓	✓	h	h	✓	✓
	ü	∗	✓	✓	✓	✓	h	h	✓
	ö	∗	✓	✓	✓	✓	h	h	✓

h = harmonic
DWs:
✓ = does exist
∗ = does not exist

What is striking is that there are no combinations such as *u–ı*, *o–ı*, *i–ı* etc.[11] Since the vowel *ı*, the realisation of an empty nucleus, does not occur after a vowel containing **I** or **U**, we have to assume that *ı* is always

[9] This is also implicit in Clements – Sezer (1982).

[10] Thanks to Pınar Kuğu, İstanbul, for her help in compiling a list of ∼2000 disharmonic words. For further investigations into combinatorial possibilities cf. Clements – Sezer (1982); Kabak et al. (2008). Note that Clements and Sezer (*op.cit.*) have fewer combinations than given here.

[11] Note that the chart in (5) probably contains too many combinations, i.e., it contains words that might not even be fully assimilated yet and where it is questionable whether they should be counted in at all.

harmonised. If there is an **I** or **U** preceding it, this **I**/**U** will always spread. This insight will be the basis of an alternative analysis of DWs.

To complete our picture of DWs, let us finally have a look at very recent loans such as *blendır* 'blender', *ketıl* 'kettle', *enkır* 'anchorman', *hemstır* 'hamster', *printır* 'printer' etc. Sometimes such words are regularised by older speakers (*blendır* 'blender' → *blendir*), and clearly assessed as weird and/or non-Turkish even by the younger generation. Given their questionable status, i.e., whether they are treated as real Turkish words, I will not consider them as real counterexamples to my statement about the distribution of *ı*.

4. An alternative proposal

In the last section we considered the dummy morphology approach to DWs, which makes rather problematic predictions and which I therefore reject.[12] What would an alternative account look like? Instead of assuming that DWs are compounds, I will make three claims, two of which are given in (6). The third claim, on the behaviour of **U**, will be dealt with in section **5**.

(6) (C1) Non-initial positions can contain any PE.

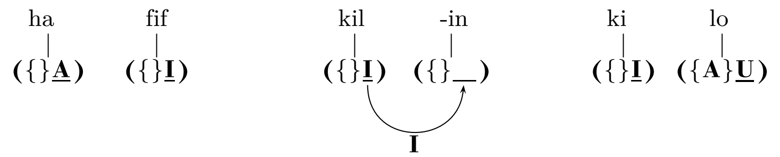
(C2) **I** only spreads into empty-headed positions.

[12] Polgárdi (1998), steering clear of dummy morphology, argues that most exceptions can be explained as a derived environment effect: In *kitap* 'book', the *a* is not harmonised as it is part of the stem and no (morphological) derivation takes place. (This of course requires that harmonic stems are harmonic to begin with.) (Most) suffixes, on the other hand, are harmonic because they involve a morphological derivation. Space restrictions preclude a full discussion, let me just note three problems. Firstly, any notion of a derived environment effect is inexpressible in GP as it violates the Minimality Hypothesis which states that processes apply **whenever** their conditions are met, and not only in some cases (i.e., a derived environment). This contradiction is completely ignored by Polgárdi. Secondly, Polgárdi's analysis completely misses the fact that there are more vowels than generally assumed; a fact that my analysis does capture, cf. below. Thirdly, Polgárdi analyses disharmonic suffixes like *-ki* as forming one single domain with the root, thus being able to be disharmonic. However, *-ki* can be separated from the root by any number of (harmonising) suffixes, e.g., the locative *-da* in *bira-da-ki* 'the one in the beer', thus making the claim of a close connection with the root and rootlike behaviour problematic.

(C1) immediately explains why we find *hafif* 'light', *arzu* 'wish', etc. If non-initial positions are not restricted in what melody they can contain, any vowel should be possible in principle, (7a) is perfectly well-formed. (We will see in a moment how the distribution of *ı* is restricted by (C2).) (C1) makes dummy compound morphology unnecessary and thus avoids (P1) and (P2).

(C2) explains why empty nuclei are always harmonised, why *ı* never disrespects harmony. It will make sure that (P3) is avoided. An *ı* is the realisation of an empty nucleus (empty-headed by definition), and if there is an **I** in the harmonic head, it will spread into that empty position by (C2). In a word like *kilo* 'kilo' the **I** in *i* $(\{\}\underline{\mathbf{I}})$ cannot spread because N_2 contains *o* $(\{\mathbf{A}\}\underline{\mathbf{U}})$, a headed expression. Again, there is no need to resort to dummy compound structure. In (7c), **I** fails to spread not because of a compound boundary, but because the target position is not empty-headed. (C2) allows spreading of **I** into empty-headed positions only.

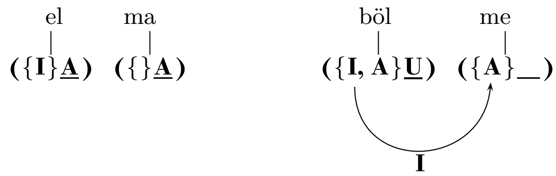
(7) (a) hafif 'light' (= (4b)) (b) kil-in 'clay-2sg' (c) kilo 'kilo'



(C1) and (C2) taken together explain most of the details of TVH. Furthermore, (C2) makes an important prediction about *a*. In DWs like *bira* 'beer' or *elma* 'apple', **I** does **not** spread from N_1 to N_2. If (C2) is correct, we have to conclude that the final *a* in both words must be **headed**. Compare this to the (short) infinitive/gerund marker *-ma* which undergoes TVH, e.g., *böl-me* 'splitting' vs. *al-ma* 'taking'. That is, in *böl-me* **I** **does** spread from N_1 (in the stem) to N_2 (in the suffix). The conclusion is that the underlying *a* of the suffix here must be **unheaded**. That is, (C2) predicts the existence of (at least) two *a*s. Crucially, the *a* in *elma* and the *a* at the end of *al-ma* 'taking' are phonetically **different**, a detail that is usually not mentioned in accounts of TVH or Turkish in general.[13] The *a* in *elma* is $(\{\}\underline{\mathbf{A}})$, i.e., headed, while the one in the suffix *-ma* is $(\{\mathbf{A}\}\_)$, i.e., unheaded. As a consequence, the empty nucleus $(\{\}\_)$ forms a natural class with $(\{\mathbf{A}\}\_)$: both are headless and thus, both undergo **I**-spreading by (C2).

[13] For details on how analyses of Turkish phonology lump together various phonetic qualities cf. e.g., Kılıç (2003).

(8) (a) elma 'apple' (b) böl-me 'splitting'



In this context, let us note also that many, though certainly not all, DWs contain long vowels in the N_2-position, e.g., *dünya* 'world' *ceza* 'punishment' *ebru* 'marbling', etc. The final vowel in all these words is long and therefore has to be headed by a general assumption of GP that only headed expressions can be long (Ploch 1995, 95). That is, the theory requires that those vowels be headed, which in turn predicts that they should be immune to TVH. This is correct.

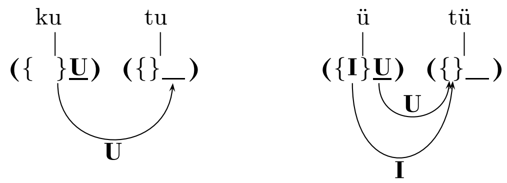
5. Extending the analysis to U-harmony

Let us now come to **U**-harmony, which will be taken care of by my third claim, given in (9).

(9) (C3) The element **U** spreads into empty expressions only.

Let us take two examples, *kutu* 'box' and *ütü* 'iron'. In both words, the second nucleus can be considered as underlyingly empty. Any melodic content it has comes from the left.

(10) (a) kutu 'box' (b) ütü 'iron'



In (10a), the first nucleus contains **U** (and nothing else) and this **U** will spread to the right since the condition in (C3) is met: the target position is completely empty. We derive *kutu*. In (10b), the first nucleus contains both **I** and **U**. According to (C3), **U** will spread if the target is completely empty, which is the case. According to (C2), **I** will spread if the target position is empty-headed and **({}_)** is empty-headed. Both **U** and **I** spread and we derive *ütü*.

The discussion of (10b) raises several questions. First of all, there seems to be an issue of ordering: If **U** is rather restricted in its spreading and only spreads to positions that are completely devoid of melodic material, then is it not the case that we need to take care that **U** spreads before **I**? After all, if **I** could spread before **U**, then the target would no longer be empty and spreading of **U** would be blocked, would it not? Note that such a claim (on who spreads first) would be inexpressible in GP as it brings us into trouble with the MH given in (1). The MH not only excludes exceptions but also ordering: If processes take place whenever they can, then there is no way of holding **I**-spreading at bay until **U** has had its chance to spread.

This problem stems from a particular conception of spreading. We usually conceive of it as an operation whereby some melodic property that initially was only present in position x ends up being present in position x as well as position y. If **I** spreads from position x to position y under such a view, then position y would not seem to be empty any more, thus creating the problem with ordering. As a matter of fact, such a conception of spreading is not the only one possible. Alternatively, spreading can also be seen as a matter of interpretation only: "Spreading" from x to y really only means that y receives the **same** interpretation as x. There is no implication that a particular element is present in both positions. This is the understanding proposed in Harris (1994, 164ff) and Charette (2007).[14] The implications of this for our current problem are clear. The final nucleus in (10b) does not contain any melody by itself, anything we hear in that position is a matter of interpretation only: **I** and **U** spread from the initial nucleus to the final nucleus, but this does not change the fact that the final nucleus is empty. Since this does not change, there is simply no problem of ordering.

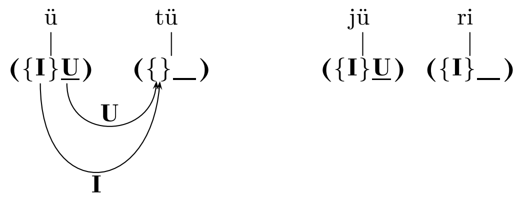
The second issue that needs to be addressed with respect to (10b) is the question of headedness. This in turn will also make clear why (C3) requires that the target nucleus be entirely empty in order for **U**

[14] It is also the understanding employed in GP 2.0, where the ongoing reduction of melodic primes often makes it impossible to spread melody, as there are many cases where there is simply no melody to spread. A notion of positions being interpreted alike without a given element being present in both of those positions is inevitable. In order to avoid the unwanted connotations that the term "spreading" carries in phonology, i.e., the physical presence of a melodic property in two positions, GP 2.0 uses the term *melodic command* (*m-command* for short), which is more neutral. For further discussion of the difference between "real" spreading as opposed to mere interpretation cf. Pöchtrager (2006, 80–4).

to spread, while (C2) allows **I** to spread as long as the head position is empty (which does not imply that the operator position has to be empty, too.) Charette and Göksel (1996) claim that when **U** spreads, it spreads as a head: It has to be a head in its start position and will also be the head in the position where it lands. This is clearly not correct. Final short high vowels (i.e., those without **A**) have to be lax (= headless) by a general phonological process in Turkish (Göksel–Kerslake 2005), i.e., in a word like *kutu* 'box' the first *u* is tense (= headed), i.e., **({}U̲)**, while the second one is lax (= headless), i.e., **({U}__)**, since it is at the end of the word and is short.[15]

This has consequences for TVH; consider (11).

(11) (a) ütü 'iron' (= (10b)) (b) jüri 'jury'



We have already discussed how in (11a/10b) **I** and **U** spread into the empty nucleus. The target nucleus is in final position and short, so neither one will be interpreted as head; the resulting interpretation will be like (lax) **({I, U}__)**. (11b) is slightly different in that the second nucleus contains melody underlyingly, viz., the element **I**. Again, we are dealing with a short nucleus in final position, there is no **A** and the expression has to be headless: **({I}__)**. Notice that in (11b), in contrast to (11a), **U** fails to spread even though the target nucleus is headless (as per phonological laxing). This justifies (C3): **U** is not content with target positions that are empty-headed; rather, they have to be completely empty.[16]

[15] This is not the complete story of final laxing. Also final vowels containing **A** differ somewhat from their word-internal counterparts: the *a* in the (harmonising) ablative suffix *-dan* differs from the *a* in the (harmonising) locative suffix *-da*, which in turn differs from the (non-harmonising) *a* at the end of *elma* 'apple'. In other words, there would not only be a second type of *a* as suggested here, but also a third one. This would require two different unheaded expressions containing only **A**, impossible to express in the standard theory of GP. For discussion of the general problems surrounding the element **A**, cf. Pöchtrager (2009).

[16] Obviously, the final *i* in *jüri* 'jury' in (11a) could not have received its **I** from the first nucleus: If it had, its position would have been empty to begin with and

Our discussion has shown that **U** is more "picky" than **I**. This is captured by (C3) expressing a stronger restriction than (C2). This asymmetry has long been noted, of course, and is also amply discussed in Charette–Göksel (1996). Their solution is to claim that **U** can only spread as a head, in order to prevent a suffix like locative *-da* (where the nucleus contains a headed expression for them) from changing into **-do*. As we have seen, this is inadequate for two reasons: Firstly, the discussion of *kutu* 'box' and *ütü* 'iron' has shown that **U** can be a non-head. Secondly, an alternating suffix like *-da* is claimed to have **(**{}**A̲)** as its nucleus by Charette and Göksel; presumably they would posit the same nuclear expression in the second nucleus of *kitap* 'book', which, of course, makes the distinction between the two kinds of *a* impossible. The solution proposed in (C3) not only explains why **U** fails to spread into *-da*, but also works for DWs, which Charette and Göksel (1996) do not deal with.[17]

6. Morpheme structure: some speculations

Let us come back to (C1) in (6): I propose that non-initial positions are unrestricted in what melody they contain. If that is the case, then why are there hardly any native DWs? In principle, any combination (with the exception of *ı*, of course) should be possible.

Firstly, polysyllabic words often involve suffixes, cf. (2). Turkish suffixes, I submit, are indeed melodically restricted, i.e., I agree with Charette and Göksel (1996) in this respect. The majority of them really only contain nuclear **(**{}**__)** or, counter to Charette–Göksel (1996), are **headless (**{**A**}**__)** in their lexical representation. In other words, while (C1) allows for quite some freedom, this does not mean that there could not be additional restrictions. (C1) has to be qualified, and the qualification has to do with the morphological status of a position.[18]

therefore we would expect both **I** *and* **U** to spread from the *ü* (**jürü*), counter to fact. Note that this also predicts the absence of sequences like *e/i* plus final short *u*, since in such cases the final short *u* would be lax (unheaded) and hence a target for **I**-spreading (yielding *ü*). So far I have not found any counterexample to this prediction; the *u* in *ebru* 'marbling' etc. is long.

[17] Except for footnote 6 on pp. 20–21, where they appeal to dummy morphology.

[18] Such a restriction in affixal material is in no way unusual: English suffixes that consist of one consonant only are exclusively coronal (i.e., contain the element **A**): *-ed*, *-s*, *-th*, *-en*.

Note furthermore that even suffixes sometimes make use of the freedom given by (C1): *-(y)ken* 'while doing' never alternates, hence contains *e* (the elements **A** and **I**) lexically, e.g., *bakar-ken* 'while looking'. The relativiser *-ki* has two shapes underlyingly: In one, the nucleus contains **I** lexically, in which case there is no alternation, e.g., *dün-ki* 'yesterday's'. The other version has an empty nucleus, which undergoes TVH like any other empty nucleus: *dün-kü* 'yesterday's'.[19]

Secondly, a huge number of bisyllabic harmonic words ends in *-k*. Denwood (2002, 166) gives *sokak* 'street', *bozuk* 'broken', *börek* 'pastry', *höyük* 'mound', *kulak* 'ear', *buruk* 'twisted', *sıcak* 'hot', *ışık* 'light', *inek* 'cow', *silik* 'worn out', *ayak* 'foot', *balık* 'fish', *bebek* 'baby', *beşik* 'cradle', *kürek* 'shovel', *büyük* 'big', but the examples could easily be multiplied. This suggests that the forms were historically complex, even if not synchronically so, cf. also Muratov (1975); Denwood (1998). But if all those words contained a suffix at some point and if the second nucleus of all such words was part of the suffix and if it is furthermore correct, as I suggest, that nuclear expressions of suffixes are indeed restricted to **({}__)** or **({A}__)**, then the scarcity of native disharmonic words follows. If the vast majority of polysyllabic native words is indeed historically complex, then disharmony never had much of a chance to occur. Native stems that are polysyllabic and **not** suspicious of (historical) complexity are rare, if not inexistent (Nakipoğlu–Üntak 2008). Many polysyllabic stems are loans from languages without harmony and, as such, show the lack of harmony that characterised the donor languages, e.g., *kitap* 'book' from Arabic.

7. Outlook and conclusion

As I hope to have shown, Turkish has more vowels than assumed so far; i.e., nuclear expressions in the language are not as restricted as was claimed in Charette–Göksel (1996). The next question then will obviously be: What restricts the set of nuclear expressions we find? Charette and Göksel (1996) proposed licensing constraints in order to answer this question; but it remains to be seen how this will be done in the further

[19] The literature sometimes treats this as "optional" application of TVH, cf. Göksel–Kerslake (2005, 25). Optionality of a process with respect to a given morpheme is of course not expressible in GP; the only viable option is to posit two alternative underlying forms as is done here.

development of GP that is currently underway, viz., GP 2.0 (Pöchtrager 2006; 2009; Živanović–Pöchtrager 2010). It might well turn out that licensing constraints can be or have to be replaced by more adequate means, and that of course will also have a bearing on the analysis of Turkish.

Furthermore, the chart in (5) includes very rare combinations/those that are sometimes modified (without becoming more harmonic, though). While this should not affect the validity of the claims discussed here, there could be further restrictions, which I will leave for further research to reveal. Despite those open questions, what we have achieved is a unified treatment of harmonic and so-called disharmonic words. In fact, the claim goes deeper: There are simply no DWs (i.e., no exceptions). DWs are an artefact of a faulty analysis of TVH. Once TVH is given a proper characterisation, the DWs disappear. The exception proves the (original) rule—wrong.

References

Charette, Monik 2007. Turkish domains. In: Meltem Kelepir – Balkız Oztürk (eds): Proceedings of WAFL 2: Workshop on Altaic Formal Linguistics (MITWPL 54), 1–20. MIT Press, Cambridge MA.

Charette, Monik – Aslı Göksel 1994. Vowel harmony and switching in Turkic languages. In: SOAS Working Papers in Linguistics & Phonetics 4 : 29–56.

Charette, Monik – Aslı Göksel 1996. Licensing constraints and vowel harmony in Turkic languages. In: SOAS Working Papers in Linguistics & Phonetics 6 : 1–25.

Clements, George N. – Engin Sezer 1982. Vowel and consonant disharmony in Turkish. In: Harry van der Hulst – Norval Smith (eds): The structure of phonological representations, part II, 213–55. Foris, Dordrecht, Cinnaminson.

Denwood, Ann 1998. A template for Turkish. In: SOAS Working Papers in Lingusitics and Phonetics 177–90.

Denwood, Ann 2002. Vowel harmony in Government Phonology: Three for the price of one. In: Altai Hakpo. Journal of the Altaic Society of Korea 12 : 155–82.

Göksel, Aslı – Celia Kerslake 2005. Turkish: A comprehensive grammar. Routledge, London.

Harris, John 1994. English sound structure. Blackwell, Oxford & Cambridge MA.

Kabak, Barış – Eva Kasselkus – Kazumi Maniwa – Silke Weber 2008. Vowel harmony has direction and context: A corpus study. Paper presented at the 16th Manchester Phonology Meeting, 22–4 May 2008.

Kardela, Henryk – Bogdan Szymanek (eds) 1996. A Festschrift for Edmund Gussmann from his friends and colleagues. The University Press of the Catholic University of Lublin, Lublin.

Kaye, Jonathan 1990. 'Coda' licensing. In: Phonology 7 : 301–30.

Kaye, Jonathan 1992. Do you believe in magic? The story of s+C sequences. In: SOAS Working Papers in Linguistics & Phonetics 2 : 293–313. Reprinted in Kardela – Szymanek (1996, 155–76).

Kaye, Jonathan 1995. Derivations and interfaces. In: Jacques Durand – Francis Katamba (eds): Frontiers of phonology: Atoms, structures, derivations, 289–332. Longman, Harlow.

Kaye, Jonathan 2001. Working with licensing constraints. In: Katarzyna Dziubalska-Kołaczyk (ed.): Constraints and preferences, 251–68. Mouton de Gruyter, Berlin & New York.

Kaye, Jonathan – Jean Lowenstamm – Jean-Roger Vergnaud 1985. The internal structure of phonological representations: A theory of charm and government. In: Phonology Yearbook 2 : 305–28.

Kaye, Jonathan – Jean Lowenstamm – Jean-Roger Vergnaud 1990. Constituent structure and government in phonology. In: Phonology 7 : 193–231.

Kılıç, Mehmet Akif 2003. Türkiye türkçesi'ndeki ünlülerin sesbilgisel özellikleri [The phonetic properties of Turkish vowels]. In: A. Sumru Özsoy – Didar Akar – Mine Nakipoğlu-Demiralp – E. Eser Erguvanlı-Taylan – Ayhan Aksu-Koç (eds): Studies in Turkish linguistics. Proceedings of the Tenth International Conference in Turkish Linguistics, 3–18. Boğaziçi University Press, İstanbul.

Muratov, S. N. 1975. Some regular changes of the phonetic structure of root in the Altaic languages. In: Louis Ligeti (ed.): Researches in Altaic languages, 173–9. Akadémiai Kiadó, Budapest.

Nakipoğlu, Mine – Aslı Üntak 2008. A complete verb lexicon of Turkish based on morphemic analysis. In: Turkic Languages 12 : 221–80.

Ploch, Stefan 1995. French nasal vowels – A first approach. In: SOAS Working Papers in Lingusitics and Phonetics 5 : 91–106.

Polgárdi, Krisztina 1998. Vowel harmony. An account in terms of government and optimality. Holland Academic Graphics, The Hague.

Pöchtrager, Markus Alexander 2006. The structure of length. Doctoral dissertation, Universität Wien.

Pöchtrager, Markus Alexander 2009. Syntaxy Government Phonology. Paper presented at Generative Approaches to Contrastive Linguistics 3, University of Cyprus, Nicosia, 15–16 May 2009.

Živanović, Sašo – Markus Alexander Pöchtrager 2010. GP 2.0 and Putonghua too. In: Acta Linguistica Hungarica 57 : 357–80.