L1 transfer in L2 learning:
compound forms in the speech of Turkish learners of Greek

Marina Tzakosta
University of Crete & Greek State Scholarships Foundation
martzak@edc.uoc.gr

Abstract: This paper investigates the capacities of Turkish learners of Greek to form one-word compounds. Using an off-line experimental task we tested the mechanisms involved in the formation of existing and novel compounds in Greek. The results revealed that Turkish learners of Greek draw from the same pool of repair mechanisms as native speakers do in order to form compounds. The findings further highlighted that word formation accuracy is attributed to the level of language proficiency as well as the typological relevance between the learners’ L1 and the target L2 language.

Key words: compounding, word formation mechanisms, second language learning, language processing, language proficiency, transfer.

1. Introduction

Compounding is a word formation process which reveals the underlying productive capacities of native speakers to compose new words. According to Selkirk (1982) a compound is a word structure consisted of two major constituents each one belonging to the category of either a noun (N), an adjective (A), a verb (V) or a preposition (P). Nominal compounds tend to be more frequently attested cross-linguistically (Becker 1992, for German, Booij 1992, for Dutch, Kiefer 1992, for Hungarian, Anastasiadi-Simeonidi 1983, Anastasiadi-Symeoni 1986, 1996a, 1996b, Ralli 1992, 2002, 2003a, b, 2005, for Greek). Another global characteristic of compounds is that their heads emerge at the right edge of the word due to the Right-Hand Head Rule (Williams 1981); according to this rule, the rightmost compound constituent determines the compound’s grammatical category or semantic meaning. This is the reason why the rightmost constituent is considered to be the compound head.

Specifically for Greek, it has been claimed that compounds belong to two major categories; lexical and morphosyntactic. Lexical compounds are formed at the level of the lexicon while morphosyntactic compounds are the product of the interaction of morphology with syntax (Drachman & Malikouti-Drachman 1994). Drachman & Malikouti-Drachman (1994) propose three compound types, namely [stem + stem], [stem + word] and [word + word] words. The first two types are lexical compounds resulting from the activation of word formation processes; the latter type mirrors morphosyntactic forms which are the product of processes activated at the phrase level. Most researchers working on Greek compounding accept the same compound distinction schematized in (1) below (cf. Malikouti-Drachman 1995, Nespor & Ralli 1994, 1996). It is interesting that the main characteristic differentiating the compound structures in (1a) and (1b) is stress assignment. In the first case, the compound form is the product of multiple word formation processes, i.e. stem blending, derivation and inflection; therefore, stress is assigned by default to the unpenultimate syllable (cf. Revithiadou 1999). In the second case, the second constituent, being used as a full

1 [stem + stem] is another way of referring to [[stem + stem] + deriv. suffix] forms.
word, preserves its lexical stress. Finally, in (1c), the compound acts as a phrase, therefore, both constituents retain their accentual properties intact.

(1)  a. [stem + stem]  
     pali-ό-filos  
     ‘old friend- MASC.SG.NOM.’  

b. [stem + word]  
     pali-o-filos  
     ‘old friend-MASC.SG.NOM.’  

c. [word + word]  
     pedi thávma  
     ‘wonder child-NEUT.NOM.SG.’

Revithiadou (1995) suggests that the shape of Greek compounds is determined by phonological properties. According to her proposal, the prosodic shape of [stem + word] compounds determines their surface realization. In other words, stress shifts to the third syllable from the end when the compound constituents are part of a prosodic word. This is illustrated in (2a). In case the relationship between compound constituents is not tight, i.e. when none of the compound constituents belongs to the same prosodic word, the second constituent retains its accentual characteristics, as shown in (2b).

(2)  a. [Stem + word] PrWd  
     lemon-ό-dasos  
     ‘lemon forest-NEUT.NOM.SG.’  

b. stem + word  
     lemon-o-dáos  
     ‘lemon forest-NEUT.NOM.SG.’

The aim of the present study is to evaluate the theoretical claims posed for Greek (cf. Pάλλη 1999, Nespor & Ralli 1996, Ralli 2005, among others) by testing the capacity of Greek native speakers and Turkish learners to form existing and non-existing/ novel compound words. We place emphasis on (a) the compounds’ internal structure, (b) issues related to compound headedness, (c) the relation holding between the compound elements and (d) the status of the linking vowel. In the next section we present the methodology followed in the design of the experimental task used in our research.

2. Methodology

For the purposes of this study we designed an off-line task which took the form of two questionnaires that subjects had to fill in with written forms (named Test 1 and Test 2, abbreviated T1 and T2). Subjects had to form 215 existing compounds in T1 and 175 novel forms in T2. All grammatical categories were tested, nominal and verbal. In Test 1, subjects had to form existing/ real compound words while they had to form non-existing/ novel compounds in Test 2. Existing compounds are forms which respect word formation rules but are also semantically acceptable. Novel compounds, on the other hand, satisfy word formation rules but are semantically ambiguous or vague. To give an example, the data in (3a) is a perfectly acceptable compound in Greek both at the morphological and semantic level. (3b), on the other hand, is well-formed only at the morphological level; however, it is a semantically non-acceptable word since ‘needle forests’ do not exist. In both cases, the two constituents are joined with the linking vowel –o- which also receives stress resulting in the unmarked [stem + stem] type.2

2 We consider [stem + stem] forms to be the unmarked compound type because stress is assigned on the unpenultimate syllable (cf. Revithadou 1999).
Participants were asked to form compounds by giving answers to questions like “how is a pie made of spinach called” or “what is the name of a set of a fork and a knife”. Both tests were distributed to all subjects participating in the experiment. Subjects were categorized into three groups. One group of 40 native speakers of Greek (G1) who ranged in age between 18-58 years and all held a High School diploma or a University degree. The native speakers’ group also served as a control group. The next two groups consisted of 20 Turkish L2 Learners of Greek who ranged in age between 10-12 years. One group (G2) consisted of 10 4th grade pupils and the other group (G3) consisted of 10 6th grade pupils of Primary Education. We decided to have two distinct L2 learners’ groups because their age and class level are indicative of their L2 proficiency level. It is worth mentioning that the participants felt exceptionally awkward when they had to form non-existing compounds because of the semantic vagueness of the latter. Moreover, participants took at least double time to fill in T2 compared to T1. The mean time of filling in both tests was a quarter of an hour for T1 and 45 minutes for T2.

3. Performance of Greek native speakers
In this section, the discussion of the Greek findings is based on the four axes that our experimental design relied on; a) the compound types more frequently used by Greek native speakers, b) the landing site of heads and the degree of their accurate realization at both a perceptual and a production level, c) the role of the linking vowel and d) the emergence of variable forms being the product of the relationship between compound constituents (e.g. exocentric vs. endocentric).

The data reveal that native speakers prefer [stem + stem] compounds to [stem + word] forms in T1. [stem + stem] forms appear in 95% of the attested cases, while [stem + word] compounds in 5% (data in 4a). In T2, Greek subjects exhibit equivalent results; [stem + stem] compounds appear in 90% of the tested cases while only 10% of [stem + word] forms are produced (data in 4b).

(4a) pefk-ό-dasos ‘pine forest-NEUT.SG.NOM.’ (T1)
(4b) velon-ό-dasos ‘needle forest-NEUT.SG.NOM.’ (T2)

Our proposal is that the preference for [stem + stem] over [stem + word] forms is determined by phonological principles (following Revithiadou 1995). As illustrated in figures 1 and 2, respectively, [stem + stem] forms correspond to prosodic words which are more coherent compared to the ones mirrored by [stem + word] ones. Our claim is

---

3 Given the form questions take in the tests, it is inevitable that answers may be biased in the sense that subjects may order compound constituents based on the order in which constituents are given in the questions. Such a bias is hard to avoid even if the off-line task takes the shape of a picture naming task. However, it is worth noting that such a deviation is attested in the ‘variable forms’ category. We assume that certain compound categories allow such deviations while others do not. This issue is amenable to future research.
4 We decided to recruit adult native speakers rather than children because we assume that adults have accurate and complete knowledge of their mother language. It would be very interesting, though, to investigate the degree of deviation in the results of preschool and primary school children as opposed to those of Greek native adults. This topic is left for future research.
5 For a more detailed discussion the reader is referred to Tzakosta (2009a).
6 See Revithiadou (1995) for an equivalent account based on phonological grounds.
that the least a prosodic word branches the more coherent it is. This would explain why
\[\text{stem + stem}\] compounds are more prevalent than \[\text{stem + word}\] ones.

The linking vowel \(-\text{o}\)- always appears in compounds whose second member starts with
a consonant in both T1 and T2. Representative examples are provided in (5a-b). \(-\text{o}\)-
always emerges in \[\text{stem + stem}\] words when it is stressed, as displayed in (5c-d). The
linking vowel appears in environments where it is expected to be phonologically and
morphologically prohibited, i.e. in vocalic \(V + V\) sequences as well as \[\text{stem + word}\]
compounds. Such data are semantically indifferent as shown in the examples in (5e-f)).

\begin{itemize}
\item (5a) xion-\(\text{o}\)-nero/ xion-\(\text{o}\)-vroxo vs. xion\(\Theta\)ántropos
\hspace{1cm} \text{‘sleet-NEUT.SG.NOM.’} \hspace{1cm} \text{‘snowman-NEUT.SG.NOM.’} \hspace{1cm} \text{(T1)}
\item (5b) vrox-\(\text{o}\)-vrodí vs. drak\(\Theta\)ántropos
\hspace{1cm} \text{‘rain & thunder-FEM.SG.NOM.’} \hspace{1cm} \text{‘dragonman-MASC.SG.NOM.’} \hspace{1cm} \text{(T2)}
\item (5c) kocin-\(\text{o}\)-aspros
\hspace{1cm} \text{‘red & white-MASC.SG.NOM.’} \hspace{1cm} \text{(T1)}
\item (5d) mis-\(\text{o}\)-ilios
\hspace{1cm} \text{‘half sun-MASC.SG.NOM.’} \hspace{1cm} \text{(T2)}
\item (5e) vori-\(\text{o}\)-anatolikos vs. vori\(\Theta\)anatolikos
\hspace{1cm} \text{‘north-east-ADJ.MASC.SG.NOM.’} \hspace{1cm} \text{(T1)}
\item (5f) ner\(\Theta\)éboros vs. ner-\(\text{o}\)-éboros
\hspace{1cm} \text{‘water seller-MASC.SG.NOM.’} \hspace{1cm} \text{(T2)}
\end{itemize}

Finally, heads are almost across-the-board accurately perceived in T1. There is only one
case of a native speaker who, in T1, forms the compound \text{dasolémono}, ‘forest lemon –
\text{NEUT.NOM.SG.}’ instead of the correct \text{lemonódasos}, ‘lemon forest \text{NEUT.NOM.SG.}’.
On the other hand, compound heads are ambiguously perceived in circa 5% of the cases
in T2. Some relevant data are given in (6).

\begin{itemize}
\item (6a) spanak-\(\text{o}\)-rizo \text{‘spinach rice - NEUT.SG.NOM.’} \hspace{1cm} \text{(T1)}
\item (6b) kreat-\(\text{o}\)-rizo vs. *riz-\(\text{o}\)-kreas \text{‘meat rice/ *rice meat - NEUT.SG.NOM.’}
\item (6c) zo-\(\text{o}\)-dromos vs. *drom-\(\text{o}\)-zoos
\hspace{1cm} \text{‘animal road/ *road animal - MASC.SG.NOM.’} \hspace{1cm} \text{(T2)}
\end{itemize}

Our basic assumption is that head misperception is attributed to ambiguous or vague
semantic synthesis of novel compounds. In addition, it is important to note that
subordinate compounds exhibit a high rate of variation but no formation errors are
attested. This is exemplified in (7) where stressed syllables and inflectional endings are
bolded for the ease of reading and recognizing different compound types. Native
speakers seem to recognize the possibility for variation in both tests without any ‘cost’ in meaning. As in non-variable forms, the preferred compound type of variable forms is [stem + stem].

(7a) xtip-o-kárdi vs. Kardi-ó-xtipos/ kardi-o-xtípi
‘heart beat - MASC./ NEUT.SG.NOM’ (T1)
(7b) xtip-o-kéfalo vs. kefal-ó-xtipos/ kefal-o-xtípi
‘head clack - MASC./ NEUT.SG.NOM.’ (T2)

In sum, the Greek data reveal that the preferred compound type is the [stem + stem] type. The linking vowel emerges in environments where its presence is phonologically driven or cases in which it does not cause any difference in meaning. Existing compounds are right headed, while novel compounds display some extent of variation because of the lack of semantic transparency of the newly formed words. In the next section, the data from Turkish L2 learners of Greek will be discussed on the basis of the same properties.

4. Compounding in Turkish

The major characteristic of Turkish compounding is that compounds appear as two-word forms, as shown in (8a) (Ζεγκίνης & Χιδίρογλου 1995, Kornfilt 1997). This means that the shape of Turkish compounds is closer to that of Greek ‘loose’ morphosyntactic compounds rather than lexical one-word compounds. According to Kornfilt (1997), there are three compound types in Turkish; bare nouns, adjective nouns and (s)I compounds. Bare nouns consist of two joined elements without having undergone suffixation, as exemplified in (8b). Adjective nouns consist of two single words, an adjective and a noun, as in (8c), while s(I) compounds are the most common ones (8d) (Kornfilt 1997, Ζεγκίνης & Χιδίρογλου 1995). In s(I) compounds the –I element is a semantically vacuous element placed at the right edge of the word, acting as a linking element. Linking vowels [e/a] appear in internal position only in verbal compounds.

(8a) Kar-dan ‘snow-ABL.’ > kardan adan ‘snowman’
(8b) Avukat kadin ‘woman lawyer’
(8c) Kara fatma ‘cockroach’
(8d) Ana ‘mother’ + dil ‘language’ < Anadili ‘mother language’

Given their loose synthesis, [word + word] compounds undergo word internal inflection. Compounds consist of one word in case, first, they have undergone certain phonological processes, second, they are inflected word-internally, third, they are scientific terms, fourth, they are of Persian origin, fifth, they have a modal verb as a second constituent and, finally, they are toponyms. As far as stress is concerned, most compounds are stressed on the first elements, as illustrated in (9a). Finally, heads appear at the right edge of the word as shown in (9b) (cf. Kornfilt 1997).

(9a) Bûz + dolabi (lit. ice cupboard) ‘refrigerator’
(9b) Gel-e-bilmek ‘able to come-1PRES.IND.SG.’

Both our main Turkish references (Ζεγκίνης & Χιδίρογλου 1995 and Kornfilt 1997) do not provide accurate glosses for the data we adopt from them.
4.1. Performance of Turkish L2 learners

Tables 1-5 provide the statistical results regarding the performance of Turkish learners of Greek in the two tests. Table 1 exhibits the rates of successfully answered questions regarding different compound types. Table 2 provides the frequency rates of [stem + stem] and [stem + word] compounds for the two school grades. [stem + stem] higher frequency tests in T1 as opposed to T2. It is important to note that in 10% of the total of the answered questions, (a) L2 subjects answer with either two-word forms (10a, b, c), (b) one-word forms which undergo internal inflection (10d), (c) and single word forms with stress shifted outside of the trisyllabic window – as Greek would require – (10e, g) or (e) double stress assignment (10f). Such data reveal the influence of Turkish L1 on Greek L2.

**Table 1. Rate of unsuccessful answers**

<table>
<thead>
<tr>
<th></th>
<th>Nominal</th>
<th>Verbal</th>
</tr>
</thead>
<tbody>
<tr>
<td>4th grade</td>
<td>76.27% - 44.96%</td>
<td>100% - 100%</td>
</tr>
<tr>
<td>6th grade</td>
<td>43.14% - 52.87%</td>
<td>100% - 94.44%</td>
</tr>
<tr>
<td>Total</td>
<td>59.7% - 49%</td>
<td>100% - 97.22%</td>
</tr>
</tbody>
</table>

**Table 2. Compound types**

<table>
<thead>
<tr>
<th>Compound types</th>
<th>4th grade</th>
<th>6th grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>80% &gt; 20%</td>
<td>75% &gt; 25%</td>
</tr>
<tr>
<td>Test 2</td>
<td>60% &gt; 40%</td>
<td>57% &gt; 43%</td>
</tr>
<tr>
<td>Total</td>
<td>70% &gt; 30%</td>
<td>66% &gt; 44%</td>
</tr>
</tbody>
</table>

(10a) odod-ó-vourtsa vs. vúrta-dódja (T1)  
‘toothbrush - FEM.NOM.SG.’ vs. ‘brush–FEM.NOM.SG.–tooth–NEUT.NOM.PL.’

(10b) aspr-ó-ruxa vs. áspra rúxa (T1)  
‘white clothing–NEUT.NOM.SG.’ vs. ‘white clothes’

(10c) spanak-ó-rizo vs. spanáki rízo (T1)  
‘spinach rice – NEUT.NOM.SG.’ vs. ‘spinach rice’

(10d) likOánthropos vs. likos(+)+ánthropos (T1)  
‘wolfman – MASC.NOM.SG.’ vs. ‘wolfman’

(10e) astrap-o-vrodí vs. astráp-o-vodi (T1)  
‘thunderclap – FEM.NOM.SG.’ vs. ‘thunderclap – NEUT.NOM.SG.’

(10f) aster-o-kinigós vs. Astér-o-kinigós (T2)  
‘starhunter – MASC.NOM.SG.’ vs. ‘starhunter’

---

8 In all tables two numbers are provided in all cells. The leftmost number mirrors the rates exhibited by [stem + word] compounds and the rightmost mirrors the rates exhibited by [word + word] compounds. The totals represent the mean rate for each compound type.
9 The actual orthographic form was ‘σπανάκυ ρύζο’.
Table 3 unfolds the extensive use of the linking vowel. The leftmost number shows the use of the linking vowel in phonological environments where it is necessary, the middle number displays the rate of linking vowel misuse – it does not appear in environments in which it is necessary -. Finally, the rightmost number reveals the rate of emergence of the linking vowel in environments in which it is not necessary. The use of the linking vowel is decreased in T2. The examples in (11) illustrate instances of data where the uncertainty regarding the linking vowel leads the subjects to the use of two-word forms (11a, b) or to internally inflected forms (11c, d).

<table>
<thead>
<tr>
<th>Linking vowel</th>
<th>4th grade</th>
<th>6th grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>80% &gt; 5% &gt; 15%</td>
<td>85% &gt; 3% &gt; 12%</td>
</tr>
<tr>
<td>Test 2</td>
<td>85% &gt; 7% &gt; 8%</td>
<td>45.11% &gt; 40% &gt; 14.89%</td>
</tr>
<tr>
<td>Total</td>
<td>82.5% &gt; 6% &gt; 11.5%</td>
<td>65% &gt; 7% &gt; 13.45%</td>
</tr>
</tbody>
</table>

(11a) spanak-ό-rizo vs. spanáki rízo\textsuperscript{10} (T1) ‘spinach rice – NEUT.NOM.SG.’
(11b) pag-ό-vuno vs. págo vúno\textsuperscript{11} (T1) ‘iceberg-NEUT.NOM.SG.’
(11c) pupul-ό-vuno vs. pupulávuna (T2) ‘feather mountain-NEUT.NOM.SG.’
(11d) agur-ό-pita vs. agurjápito (T2) ‘cucumber pie-FEM.NOM.SG.’

Table 4 exhibits the rates of accurate use of compound forms. Obviously, older children perform better regarding the correct position of the heads in both T1 and T2. However, (12) exemplifies cases of head misperception in both T1 and T2. In these data, actual heads appear as the leftmost compound constituents.

<table>
<thead>
<tr>
<th>Headedness</th>
<th>4th grade</th>
<th>6th grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>97%</td>
<td>98%</td>
</tr>
<tr>
<td>Test 2</td>
<td>93%</td>
<td>98%</td>
</tr>
<tr>
<td>Total</td>
<td>95%</td>
<td>98%</td>
</tr>
</tbody>
</table>

\textsuperscript{10} The actual orthographic form was ‘σπανάκυ ρύζο’. 
\textsuperscript{11} The actual orthographic form was ‘πάγο βοόνο’. 
Table 5, provides the rates of exocentric variable forms. Subjects seem to have difficulties recognizing the head constituent in variable forms. Representative examples are given in (13). In 5% of the forms which are expected to emerge as variable subjects prefer to answer with non-compound forms, as shown in (14). Tables 4 and 5 and the data in (12)-(14) underline the fact that children do not seem to be biased by the order in which compound constituents are given to them.

Table 5. Variation

<table>
<thead>
<tr>
<th>Variable forms</th>
<th>4th grade</th>
<th>6th grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>82% &gt; 18%</td>
<td>75,39% &gt; 24,70%</td>
</tr>
<tr>
<td>Test 2</td>
<td>41,17% &gt; 58,83%</td>
<td>82,08% &gt; 17,91%</td>
</tr>
<tr>
<td>Total</td>
<td>61,5% &gt; 38,5%</td>
<td>78,8% &gt; 21,2%</td>
</tr>
</tbody>
</table>

(13a) avg-o-lémono vs. lemon-o-avgó  ‘bitten egg and lemon-NEUT.NOM.SG.’
(13b) alat-o-rigani vs. alatØrigáni  ‘salt and origanum-FEM.NOM.SG.’
(14a) andr-ó-gino ‘husband & wife vs. zevgári ‘couple-NEUT.NOM.SG.’
(14b) jinek-ó-peda ‘women & children vs. ikogénia ‘family-FEM.NOM.SG.’
(14c) nixteridØánthropos ‘batman vs. vrikólakas ‘vampire-MASC.NOM.SG.’

12 The actual orthographic form was ‘δέρμω προβλήματα’.
5. General discussion

The purpose of this study was to assess the mechanisms driving compound formation in groups of native speakers and L2 learners of Greek. Our major result was that both native speakers and L2 learners of Greek draw from the same pool of learning strategies governing compounding. More specifically, Greek native speakers do not make major mistakes in the formation of existing compounds (Test 1). However, Test 2 reveals an extended degree of variation in the formation of non-existing forms. Such results highlight the fact that the formation and use of existing compound forms are massively determined by mnemonic knowledge. Consequently, the activation of word formation mechanisms is eliminated. Put differently, word formation mechanisms are clearly activated in the formation of non-existing forms (Test 2). Such activation results in variation in the produced forms. Variation is attested mostly with respect to the emergence (or not) of the linking vowel and the preferred compound structure but not issues related to headedness. In other words, heads appear as the second compound members. The linking vowel appears across-the-board in compounds whose second constituent starts with a consonant. However, its emergence is optional if the second constituent starts with a vowel. Linking vowels are always present in [stem + stem] compounds even if the second compound member starts with a vowel. We have argued that, in such cases, the use of the linking vowel is phonologically driven so that the more coherent prosodic word structure surfaces.

Turkish L2 learners of Greek, on the other hand, are highly influenced by their mother tongue. However, the recruitment of two different age groups of Turkish learners of Greek revealed that older speakers have better knowledge of Greek. More specifically, all L2 learners show a preference for [stem + stem] forms in both tests. However, the data exhibit that Greek compound formation is determined by various characteristics of Turkish compounding. Therefore, we reported a statistical prevalence of two word compounds. Moreover, compounds show internal inflection and stress shift. Regarding the emergence of the linking vowel, we notice a broad use of it. Like native speakers of Greek, Turkish L2 learners display strong activation of mnemonic knowledge in Test 1 as opposed to Test 2.

Our research program will be continued at the level of testing the validity of the above findings by comparing them with other languages. More specifically, we are interested in investigating whether L1 influence on L2 learning depends on proficiency level and/ or age as well as exploring the degree to which UG word formation principles govern L1 acquisition and L2 learning. Such findings will facilitate a complete and thorough typological analysis of compound formation in L2 cross-linguistically.

References


13 See Tzakosta (2009b, c) for a detailed investigation of the perceptual and production capacities of Dutch and German learners of Greek, respectively, regarding compound formation.
Tzakosta, M. 2009a. “Perceptual ambiguities in the formation of Greek compounds by native speakers”.