

# Βελτιωμένοι κώδικες συλλαβισμού πολυτονικῶν καὶ μονοτονικῶν νεοελληνικῶν κειμένων για τὸ T<sub>E</sub>X καὶ τὸ L<sup>A</sup>T<sub>E</sub>X

Δημήτριος Ἀ. Φιλίππου  
Κάτω Γατζέα  
385 00 Βόλος

14 Νοεμβρίου 2000

Καλοὶ μου φιλόλογοι,  
μὴ βαρᾶτε, ποὺ τόλμησα καὶ μῆκα  
στ' ἀμπελοχώραφά σας!  
Ἐνοχος ἐξ ἀμελείας

## Abstract

Existing T<sub>E</sub>X hyphenation patterns for typesetting modern Greek texts do not take into account the grammatical differences between the old polytonic (multi-accent) writing system and the new monotonic (uni-accent) writing system. Also, those patterns have been found to contain several hyphenation mistakes. To correct this, new hyphenation patterns were written from scratch based on the rules of the Greek grammar. This paper describes the logic behind the new hyphenation patterns and the inherent difficulties in automatic hyphenation of Greek texts. The new hyphenation patterns, which can be used with the `greek` option of `babel` or Dryllerakis' `GREEKTEX`, are available for free from CTAN.

## 1 Εἰσαγωγή

Ὅσοι στοιχειοθετοῦν ἑλληνικὰ κείμενα μετὰ τὸ T<sub>E</sub>X ἢ τὸ L<sup>A</sup>T<sub>E</sub>X, ἴσως δοῦν τὴν λέξη *πυκνότητα* κομμένη μεταξὺ τοῦ κάππα καὶ τοῦ νῦ, δηλ. ὡς *πυκ-νό-τη-τα*. Ἴσως ἐπίσης νὰ ἐκπλαγοῦν βλέποντας τὴν λέξη *ἰσθμός* κομμένη μεταξὺ τοῦ δευτέρου καὶ τοῦ τρίτου συμφώνου ὡς *ἰσθ-μός*. Σύμφωνα μετὰ τοὺς κανόνες τῆς ἑλληνικῆς γραμματικῆς, ἡ πρώτη λέξη συλλαβίζεται ὡς *πυ-κινό-τη-τα*, γιατί ὑπάρχει λέξη ἑλληνικὴ ποὺ νὰ ἀρχίζει ἀπὸ κάππα-νῦ, π.χ., *κινήμη*. Ὅσο για τὴν δεύτερη λέξη, ὁ σωστὸς συλλαβισμὸς τῆς εἶναι *ἰ-σθμός*, ἐπειδὴ ἓνα σύμπλεγμα τριῶν ἢ περισσοτέρων συμφώνων δὲν διασπᾶται ἐὰν ὑπάρχει λέξη ἑλληνικὴ ποὺ νὰ ἀρχίζει ἔστω ἀπὸ τὰ δύο πρῶτα σύμφωνα τοῦ συμπλέγματος, π.χ., *σθένος*.

Στο Internet βρίσκουμε τέσσερα αρχεία με κώδικες (patterns) για τὸν συλλαβισμό ἑλληνικῶν κειμένων μετὰ τὸ T<sub>E</sub>X: τὸ `rgrhyph.tex` τοῦ Γ. Χαραλάμπους [1], τὸ `grkhyphen.tex` τοῦ Κ. Δρυλλεράκη [2], τὸ `gehyphen.tex` τοῦ Γ. Μοσχοβάκη [3] καὶ τὸ `grhyph.tex` τοῦ C. Beccari [4]. Παρακάτω δὲν θὰ ἀσχοληθοῦμε μετὰ τὸ ἀρχεῖο τοῦ Δρυλλεράκη [2], ἐφ' ὅσον αὐτὸ δὲν ἀποτελεῖ τίποτα ἄλλο ἀπὸ μία περιορισμένη ἐκδοσὴ τοῦ `rgrhyph.tex` (σὲ κάποια γραμμὴ τοῦ ἀρχείου αὐτοῦ παρεμβάλλεται ἡ ἐντολὴ `\endinput` καὶ ἕνα σχόλιο ὅτι τὸ T<sub>E</sub>X παρουσιάζει πρόβλημα μνήμης). Ἐπίσης δὲν θὰ ἀσχοληθοῦμε μετὰ τὸ ἀρχεῖο κωδίκων συλλαβισμού τοῦ Μοσχοβάκη [3], ἐπειδὴ οἱ κώδικες αὐτοὶ εἶναι πολὺ περιορισμένοι καὶ ἐπειδὴ τὸ ὅλο πακέτο `greektex` τοῦ ἴδιου συγγραφέα εἶναι φτιαγμένο κυρίως γιὰ τὸ λειτουργικὸ σύστημα DOS, ποὺ τώρα πιά εἶναι ξεπερασμένο.

Ἐκτὸς ἀπὸ τὰ παραδείγματα ποὺ προαναφέραμε, οἱ κώδικες συλλαβισμού ἑλληνικῶν ὄλων αὐτῶν τῶν ἀρχείων παρουσιάζουν πολλὰ ἄλλα λάθη καὶ ἐλλείψεις. Οἱ κώδικες τοῦ Beccari [4] χωρίζουν λανθασμένα ὀρισμένα συμπλέγματα ἐκ τῶν ὁποίων ἀρχίζουν ἀρχαῖες ἑλληνικὲς λέξεις *δμ* (*δμῶς* = *δοῦλος*), *δν* (*δνοφερός* = *ζοφερός* ἢ *σκοτεινός*), *τλ* (*τλημωσύνη* = *φτώχεια* ἢ *μιζέρια*). Οἱ ἴδιοι κώδικες δὲν ἐπιτρέπουν τὸν χωρισμὸ ἀρχικῶν φωνηέντων. Τέλος, δὲν λαμβάνεται καμία μέριμνα γιὰ συμπλέγματα τριῶν ἢ περισσοτέρων συμφώνων.

Παρομοίως, στοὺς κώδικες τοῦ Χαραλάμπους [1] δὲν περιλαμβάνεται κανένα συμπλέγμα τριῶν ἢ περισσοτέρων συμφώνων, μετὰ μοναδικὴ ἐξαιρεσὴ τὸ *σθμ*, τὸ ὁποῖο παρουσιάζεται ὀρθῶς ὡς ἀδιάσπαστο. Τὰ ρινικὰ συμπλέγματα *γκ*, *μπ* καὶ *ντ* παρουσιάζονται ὡς ἀδιάσπαστα, κάτι ποὺ δὲν ἰσχύει στὴν γραμματικὴ τοῦ πολυτονικοῦ [5]. Στούς κώδικες τοῦ Χαραλάμπους βρίσκουμε ἀκόμα τὸν λανθασμένο χωρισμὸ τοῦ συμπλέγματος *τμ* (*τμήμα*) καὶ τὴν παράθεση πλήθους κωδίκων γιὰ τὴν ἀπαγόρευση (χωρὶς ἐμφανὴ λόγῳ) τοῦ συλλαβισμού τῶν δύο τελευταίων γραμμάτων κάθε λέξης.

Γιὰ νὰ διορθωθεῖ ἡ κατάστασις αὐτή, οἱ κώδικες συλλαβισμού ἑλληνικῶν γιὰ τὸ T<sub>E</sub>X κατὰ τὴν τυποποίησιν 7-bit τοῦ Levy [6] ἔπρεπε νὰ ξαναγραφοῦν. (Οἱ ἐλάχιστες διαφορὲς μεταξὺ τῶν κωδικοποιήσεων τῶν ἑλληνικῶν χαρακτήρων κατὰ Χαραλάμπους [1], Δρυλλεράκη [2] ἢ κατὰ τὴν ἑλληνικὴ ἐπιλογὴ τῆς `babel` [7] δὲν ἐπηρεάζουν τοὺς κώδικες συλλαβισμού.) Παρακάτω, θὰ δοῦμε πῶς συλλαβίζει τὸ T<sub>E</sub>X, πῶς σχεδιάσθησαν οἱ νέοι κώδικες συλλαβισμού τῶν νεοελληνικῶν (μετὰ τὶς ἰδιαιτερότητες τοῦ μονοτονικοῦ καὶ τοῦ πολυτονικοῦ), καθὼς καὶ ἕνα πείραμα ποὺ ἀποδεικνύει τὴν ἀνωτερότητα τῶν νέων κωδίκων.

## 2 Πῶς συλλαβίζει τὸ T<sub>E</sub>X

Ὁ τρόπος μετὰ τὸν ὁποῖο συλλαβίζει τὸ T<sub>E</sub>X ἕνα ὁποιοδήποτε κείμενο περιγράφεται στὸ *T<sub>E</sub>Xbook* [8]. Ὁ Χαραλάμπους [9] ὡστόσο ἐξηγεῖ καλύτερα τὸν ἀλγόριθμο συλλαβισμού τοῦ T<sub>E</sub>X:

«Ὅταν τὸ T<sub>E</sub>X δημιουργεῖ μία φόρμα ὅπως τὴν `plain` ἢ τὴν `latex`, διαβάζει ὀρισμένες πληροφορίες ἀπὸ τὸ ἀρχεῖο `hyphen.tex` (ἢ ἀπὸ τὸ `USHyphen.tex`, τὸ `FRHyphen.tex`, κ.ἄ.ἄ.), τὸ ὁποῖο περιέχει τοὺς κώδικες συλλαβισμού μίας συγκεκριμένης γλώσσας. Οἱ κώδικες αὐτοὶ

είναι σειρές γραμμάτων που χωρίζονται από αριθμούς, όπως  $x1y2z$ .  
Ἡ ἀρχὴ εἶναι ἡ ἀκόλουθη:

- Ἐὰν τὸ σύνολο τῶν κωδικῶν εἶναι κενό, τότε ὁ συλλαβισμὸς δὲν ἐπιτρέπεται.
- Ἐὰν ὑπάρχει ὁ κώδικας  $x1y$ , τότε ἐπιτρέπεται ὁ χωρισμὸς τῶν  $x-y$  σὲ ὅλες τὶς λέξεις πού περιέχουν τὸ σύμπλεγμα  $xy$ . Ἐὰν ὁ κώδικας ἔχει τὴν μορφή  $x1yzw$ , τότε ἐπιτρέπεται ὁ χωρισμὸς τῶν  $x-y$  μόνο στὴν περίπτωση ἐκείνη ὅπου μετὰ τὸ  $xy$  ἀκολουθεῖ τὸ σύμπλεγμα  $zw$ .
- Ἐὰν ὑπάρχει ὁ κώδικας  $x1y$  καὶ ὁ κώδικας  $x2yabc$ , τότε τὸ σύμπλεγμα  $xy$  χωρίζεται πάντοτε, ἐκτὸς ἀπὸ τὴν περίπτωση ὅπου μετὰ τὸ  $xy$  ἀκολουθεῖ τὸ σύμπλεγμα  $abc$ . Μὲ ἄλλα λόγια, ὁ ἀριθμὸς 2 δείχνει μίαν ἐξαιρέση στὸν κανόνα 'τὸ  $x$  χωρίζεται ἀπὸ τὸ  $y$ '.
- Τὸ ἴδιο ἰσχύει μὲ μεγαλύτερους ἀριθμούς: ὁ ἀριθμὸς 3 δείχνει μίαν ἐξαιρέση σὲ κωδικὸς συλλαβισμοῦ μὲ ἀριθμὸ 2, κ.ο.κ. [...]
- Μία τελεία ἐμπρός (ἢ μετὰ) ἀπὸ ἕναν κώδικα δείχνει ὅτι ὁ κώδικας αὐτὸς ἰσχύει στὴν ἀρχή (ἢ στὸ τέλος) μίας λέξης. Κατὰ τὸν τρόπο αὐτό, ὁ κώδικας  $.xy2z$  ἰσχύει *μόνο* στὴν περίπτωση τῆς λέξης 'xyz'.»

Ἄς ὑποθέσουμε λοιπὸν ὅτι ἔχουμε ἔχουμε μίαν σειρά ἀπὸ κωδικὸς γιὰ τὸν συλλαβισμὸ ἐλληνικῶν προσαρμοσμένους στὴν τυποποίηση 7-bit κατὰ Levy [6] (τὸ πῶς δημιουργοῦμε αὐτοὺς τοὺς κωδικὸς, θὰ τὸ ἐξηγήσουμε παρακάτω):

```
a1
e1
...
2111
...
```

καὶ ἂς ὑποθέσουμε ἐπιπλέον ὅτι θέλουμε νὰ στοιχειοθετήσουμε ἕνα κείμενο ἐλληνικὸ πού περιέχει τὴν λέξη *ἀλλά*.

Ἀρχικῶς, μὲ τοὺς κωδικὸς, ἐτοιμάζουμε μίαν κατάλληλη φόρμα, π.χ., τὴν φόρμα `mygreek.fmt`. Κατόπιν, ἐτοιμάζουμε τὸ ἀρχεῖο μὲ τὸ κείμενο πού θέλουμε νὰ στοιχειοθετήσουμε, π.χ., `mytext.tex`.

Ἐπειδὴ στὴν ἐλληνικὴ τυπογραφία ἐπιτρέπεται ἡ ἐγκατάληψη ἑνὸς καὶ μόνου χαρακτήρα στὸ δεξιὸ ἢ στὸ ἀριστερὸ ἄκρο τῆς ἀράδας, στὸ ἀρχεῖο `mytext.tex`, πρῶτα ἐπιλέγουμε τὴν γλώσσα ἐργασίας

```
\selectlanguage{greek}
```

καὶ κατόπιν ὀρίζουμε:

```
\righthyphenmin=1
\lefthyphenmin=1
```

Μετά ακολουθεῖ τὸ κείμενο

```
...  
>all'a  
...
```

Τρέχουμε κατόπιν τὸ ἀρχεῖο `mytext.tex` μὲ τὸ `TEX`:

```
tex &mygreek.fmt mytext.tex
```

Ὅταν τὸ `TEX` συναντήσει τὴν λέξη *ἀλλά* (`>all'a`), μὲ βάση τὸν κώδικα `a1` ἀποφασίζει νὰ τὴν κόψει μεταξὺ τοῦ *ά* καὶ τοῦ πρώτου *λ*. Ὅμως ὁ κώδικας `2111` ἐμποδίζει τὸ κόψιμο πρὶν ἀπὸ δύο λάμβδα. Ἐφ' ὅσον τὸ `2` εἶναι μεγαλύτερο τοῦ `1`, ὁ δεύτερος κώδικας-κανόνας ὑπερισχύει καὶ τὸ `TEX` εἶναι τελικῶς ἀναγκασμένο νὰ συλλαβίσει τὴν λέξη ὡς *άλ-λά*, ποὺ εἶναι καὶ γραμματικῶς ὀρθό.

### 3 Ἡ δημιουργία βελτιωμένων κωδίκων συλλαβισμού γιὰ τὰ νέα ἑλληνικά

Ἡ δημιουργία κωδίκων συλλαβισμού εἶναι μία ἐπίπονη προσπάθεια. Κάθε γλώσσα ἔχει τὶς ἰδιαιτερότητές της. Ἐπιπλέον, διαφορὲς στὸν τρόπο συλλαβισμού παρουσιάζονται καὶ μεταξὺ διαφορετικῶν διαλέκτων τῆς ἴδιας γλώσσας. Τὸ πλέον γνωστὸ παράδειγμα εἶναι ἡ ἀγγλικὴ γλώσσα, λέξεις τῆς ὁποίας συλλαβίζονται κατὰ ἕναν τρόπο σύμφωνα μὲ τὴν βρετανικὴ γραμματικὴ καὶ κατὰ ἄλλο τρόπο σύμφωνα μὲ τὴν ἀμερικανικὴ γραμματικὴ. Στὴν περίπτωσή τῆς ἑλληνικῆς γλώσσας, οἱ κανόνες συλλαβισμού διαφέρουν ἀναλόγως μὲ τὸ ἐὰν πρόκειται γιὰ

- ἀρχαῖα ἑλληνικά ἢ καθαρῆουσα [10],
- νέα ἑλληνικά στὸ πολυτονικὸ σύστημα τοῦ Τριανταφυλλίδη [5], ἢ
- νέα ἑλληνικά στὸ περιορισμένο πολυτονικὸ σύστημα τῆς περιόδου 1976–1981 [11] ἢ στὸ σύγχρονο μονοτονικὸ σύστημα [12].

Μία λύση γιὰ τὴν δημιουργία κωδίκων συλλαβισμού εἶναι μὲ τὴν χρῆση τοῦ προγράμματος `PATGEN` [13], ποὺ ὅμως ἀπαιτεῖ τὴν τροφοδοσίαν ἐνὸς ὀλοκλήρου λεξικοῦ σὲ ἠλεκτρονικὴ μορφή προκειμένου νὰ δώσει σωστοὺς κώδικες μὲ ὅλες τὶς πιθανὲς ἐξαιρέσεις. Μία ἄλλη λύση λιγώτερο χρονοβόρα (ἄρα πιὸ πρακτικὴ) εἶναι ἡ «ἐξέταση τῶν κρυφῶν μηχανισμῶν συλλαβισμού καὶ ἡ συγγραφή κωδίκων ποὺ ἀντιστοιχοῦν στὰ ἀναλυτικὰ βήματα [τῶν γραμματικῶν κανόνων] συλλαβισμού» [9]. Τὴν λύση αὐτή, ποὺ ἐπέλεξαν ὅλοι ὅσοι συνέταξαν κώδικες συλλαβισμού ἑλληνικῶν γιὰ τὸ `TEX`, ἀκολουθοῦμε ἐδῶ γιὰ τὴν δημιουργία νέων βελτιωμένων κωδίκων γιὰ τὸ πλήρες πολυτονικὸ σύστημα τοῦ Τριανταφυλλίδη καὶ γιὰ τὸ σύγχρονο μονοτονικὸ σύστημα. Τὸ περιορισμένο πολυτονικὸ σύστημα τῆς Μεταπολίτευσης θὰ τὸ ἀγνοήσουμε μιᾶς καὶ οἱ κανόνες συλλαβισμού τοῦ συστήματος αὐτοῦ εἶναι ὅμοιοι μ' αὐτοὺς τοῦ μονοτονικοῦ. Ἄλλωστε, εἶναι πιὸ ἐντιμὸ νὰ χρησιμοποιοῦμε τὸ πολυτονικὸ σύστημα στὴν πλήρη μορφή του καὶ ὄχι κουτσουρεμένο!

Οἱ κανόνες συλλαβισμού καὶ οἱ ἀντίστοιχοι κώδικες ἔχουν ὡς ἐξῆς:

(α) Ένα σύμφωνο ανάμεσα σε δύο φωνήεντα συλλαβίζεται με το δεύτερο φωνήεν

Με άλλα λόγια, μία απλή λέξη χωρίζεται σε συλλαβές μετά από κάθε φωνήεν: π.χ., *τη-λε-ό-ρα-ση*. Οί αντίστοιχοι κώδικες είναι: a1 e1 h1 i1 o1 u1 w1. Για τὸ πολυτονικὸ προσθέτουμε μερικούς ακόμη κώδικες για τὴν ἀποφυγὴ χωρισμοῦ μεταξὺ φωνήεντος καὶ ὑποδιαστολῆς: a4|1 h4|1 w4|1.

(β) Τὰ δίψηφα φωνήεντα δὲν χωρίζονται

Οί αντίστοιχοι κώδικες είναι: a2i a2'i ... u2<'i u2<~i. Μεταξὺ τῶν κωδίκων αὐτῶν παρεμβάλλουμε ὀρισμένες ἐξαιρέσεις: 'a3u 'e3u 'o3u 'u3i.

(γ) Οί καταχρηστικοί δίφθογγοι δὲν χωρίζονται

Αὐτὸ σημαίνει ὅτι δὲν χωρίζονται συνδυασμοί φωνηέντων ποὺ προφέρονται σὲ μία συλλαβή. Οί φθογγικοί συνδυασμοὶ τῶν καταχρηστικῶν δίφθόγγων εἶναι συνήθως τὰ *a-ι*, *ε-ι* καὶ *ο-ι* ποὺ εἴτε εἶναι ἄτονα εἴτε τονίζονται στὸν πρῶτο φθόγγο (π.χ., *ἀη-δό-νι*, *ζεῖ-μπέ-κης*, *ἀ-πόει-δα*, *ρόι-δι*, κ.ἄ.δ.), καὶ τὰ *ι-α*, *ι-ε* καὶ *ι-ο* ποὺ εἴτε εἶναι ἄτονα εἴτε τονίζονται στὸν δεύτερο φθόγγο (π.χ., *ἀ-μυα-λιά*, *θη-λιές*, *Μα-ριώ*, κ.ἄ.δ.). Οί αντίστοιχοι κώδικες συλλαβισμοῦ γιὰ τὸ TEX εἶναι: a2h a2"i ... i2a i2'a ... u2~w. Μεταξὺ αὐτῶν τῶν κανόνων παρεμβάλλονται κάποιες ἐξαιρέσεις: 'a3h ... <u3. Ὄρισμένοι συνδυασμοὶ φωνηέντων ἄλλοτε ἀποτελοῦν καταχρηστικὸ δίφθογγο καὶ ἄλλοτε ὄχι, ὅποτε προφέρονται χωριστὰ σὲ δύο συλλαβές: π.χ., *βόη-θα* καὶ *ἀ-πό-η-χος*, *ποιός* καὶ *ἦ-θο-ποι-ός*. Δυστυχῶς, τὸ TEX δὲν μπορεῖ νὰ διακρίνει τέτοια λιγώτερο ἐμφανῆ παιχνιδίσματα τῆς γλώσσας μας. Ὅμως εἶναι προτιμώτερο νὰ συλλαβίσει τὴν λέξη *κακοποιός* ὡς τρισύλλαβη (*κα-κο-ποιός*) ἀντὶ γιὰ τετρασύλλαβη (*κα-κο-ποι-ός*), παρὰ νὰ συλλαβίσει τὴν λέξη *μὰ* ὡς δισύλλαβη (*μυ-ά*) ἀντὶ γιὰ μονοσύλλαβη.

(δ) Τὰ σύμφωνα — ἀπλὰ καὶ δίψηφα — στὴν ἀρχὴ ἢ στὸ τέλος λέξεων δὲν ἀποτελοῦν συλλαβές

Οί αντίστοιχοι κώδικες εἶναι: 6b. 6g. 6gk. ... 6y. .b6 .g6 .gk6 ... .y6. Στους κώδικες αὐτούς, προσθέτουμε μερικούς ακόμη γιὰ τὴν περίπτωση ἐλλειπῶν λέξεων ποὺ τελειώνουν με ἀπόστροφο: 6'' 6b'' 6g'' ... 6y''.

(ε) Τὰ διπλὰ σύμφωνα χωρίζονται

Οί αντίστοιχοι κώδικες εἶναι: 4b1b 4g1g ... 4r1r 4>r1<r 4s1s ... 4y1y.

() Συνδυασμοὶ δύο ἢ περισσοτέρων συμφώνων, ἐκ τῶν ὁποίων δύο πρώτων δὲν ἀρχίζει ἑλληνικὴ λέξη, χωρίζονται μετὰ τὸ πρῶτο σύμφωνο

Οί αντίστοιχοι κώδικες εἶναι: 4b1z ... 4y1q. Μεταξὺ αὐτῶν τῶν κωδίκων, παρεμβάλλονται κάποιοι ἄλλοι γιὰ τὴν ἀντιμετώπιση ὀρισμένων τριπλῶν συμφωνικῶν συνδυασμῶν:

4r5g2m	(ἔρ-γμα)	tz2m	(μάνα-τζμειντ)
4r5j2m	(πορ-θμός)	sj2m	(i-σθμός)
4l5k2m	(Ἄλ-κμήνη)	4g5k2f	(Φραγ-κφούρτη)
4n5k2f	(Φραν-κφούρτη)	4r5k2m	(δαρ-κμά)
4g5x2t	(ἔ-λέγ-ξετε)	4r5x2t	(ὑ-πάρ-ξετε)
st2f	(φα-στυφουντάδικο)	4r5q2m	(δαρ-χμά)
4m5y2t	(λάμ-ψτε)		

Στους παραπάνω κώδικες για τὰ τριπλὰ σύμφωνα, θὰ μπορούσαμε νὰ προσθέσουμε καὶ ἄλλους για τὴν ἀντιμετώπιση κυρίως ξενικῶν λέξεων. Για παράδειγμα, θὰ μπορούσαμε νὰ προσθέσουμε τὸν κώδικα 4r5l2s για νὰ συλλαβισθεῖ τὸ ὄνομα *Κάλρσον* ὡς *Κάρ-λσον*. Ὅμως κάτι τέτοιο εἶναι ἐξαιρετικὰ ἐπίπονο (ὑπάρχουν πάμπολλες ξενικὲς λέξεις ποὺ περιέχουν συνδυασμοὺς τριπλῶν συμφώνων ποὺ δὲν ὑπάρχουν στὴν ἑλληνικὴ γλώσσα), ἀλλὰ ἴσως καὶ ἄσκοπο (ἴσως νὰ εἶναι πιὸ φυσικὸ νὰ χωριστεῖ ἡ λέξη τοῦ προηγούμενου παραδείγματος ὡς *Κάρλ-σον*).

(ζ) *Στὴν περίπτωση πολυτονικοῦ κειμένου, τὰ ρινικὰ συμλέγματα γκ, μπ καὶ ντ καὶ ν-τ χωρίζονται. Ἀντιθέτως, τὰ δίψηφα (ἄρρινα) γκ, μπ καὶ ντ δὲν χωρίζονται*

Αὕτῃ εἶναι ἡ μοναδικὴ ουσιαστικὴ διαφορὰ μεταξὺ τοῦ πολυτονικοῦ καὶ τοῦ μονοτονικοῦ. Στὸ περιορισμένο πολυτονικὸ [11] καὶ στὸ μονοτονικὸ [12], οἱ συνδυασμοὶ *γκ, μπ καὶ ντ* δὲν χωρίζονται ποτέ.

Ἔτσι, στὴν περίπτωση τοῦ πολυτονικοῦ καὶ μόνο, ἔχουμε τοὺς κώδικες 4g1k, 4m1p καὶ 4n1t. Τὰ *γκ, μπ καὶ ντ* εἶναι γενικῶς ἔρρινα [5]. Ὅμως, ὅταν βρίσκονται μετὰ ἀπὸ ἄλλα σύμφωνα εἶναι ἄρρινα: π.χ., *ἀρ-γκό, ἄλ-μπου-ρο, σεβ-ντάς*. Παρομοίως εἶναι ἄρρινα σὲ λέξεις ξενικῆς προέλευσης: π.χ., *κα-ρα-γκιό-ζης* (τουρκ.), *Ἴμπρα-ήμ* (ἄραβ.), *μπι-ντές* (γαλλ.), κ.ἄ.δ. Τὶς λέξεις ξενικῆς προέλευσης μετὰ ἄρρινα *γκ, μπ καὶ ντ* τὶς ἀντιμετωπίζουμε ὅπως ὁ Χαραλάμπος [9] ἀντιμετωπίζει τὶς σύνθετες λέξεις τῶν ἀρχαίων ἑλληνικῶν: Μὲ βᾶση τὸ ἐτυμολογικὸ λεξικὸ τοῦ Ἀνδριώτη [14] καταστρώνουμε ἕναν κατάλογο ἀπὸ κώδικες ποὺ ἀντιστοιχοῦν στὰ ἀρχικὰ μέρη τέτοιων λέξεων:

.giou5g2ko	γιου-γχοσλάβος
.kara5g2ki	καρα-γκιόζης
...	...
.qa5n2tr	χάντρα
.qa5n2tr	χαντρῶν

Τὶς ἄκλιτες ξενικὲς λέξεις μετὰ ἄρρινα *γκ, μπ καὶ ντ*, π.χ., *ἀ-μπα-ζούρ, τί-γκα*, κ.λπ., τὶς συγκεντρώνουμε στὸ σύνολο τῶν ἐξαιρέσεων `\hyphenation{...}`. Ὁ πλήρης κατάλογος αὐτῶν τῶν ξενικῶν λέξεων, κλιτῶν καὶ ἄκλιτων, μετὰ ἄρρινα *γκ, μπ καὶ ντ* δίνεται στὸ Παράρτημα τοῦ ἄρθρου. Εἶναι πολὺ πιθανὸ ὅτι κάποιες λέξεις ἀπουσιάζουν ἀπὸ τὸν κατάλογο. Ὅμως εἶναι πολὺ εὐκόλο νὰ προστεθοῦν σὲ μελλοντικὲς ἐκδόσεις τῶν κωδίκων συλλαβισμού.

Για τὸ μονοτονικὸ, οἱ μόνοι κώδικες ποὺ προσθέτουμε εἶναι για ὀρισμένες ἐμφανεῖς περιπτώσεις ἔρρινων τριπλῶν συνδυασμῶν: 4g1kt (ἔ-λεγ-κτής), 4m1pt (ἀ-μεμ-πτος), 4n1tz (νε-ραιν-τζιά) καὶ 4n1ts (βιο-λον-τσέ-λο).

(η) Καταλήξεις διπλών ή τριπλών συμφώνων δεν αποτελούν συλλαβές

Πρόκειται για όρισμένες αρχαϊκές καταλήξεις ή καταλήξεις ξενικών λέξεων που έχουν εισέλθει άκλιτες στην γλώσσα μας, όπως -κλ (πι-νάκλ), -νς (έλ-μινς), κ.ά.δ. Οι καταλήξεις αυτές συνήθως αποτελούνται από δύο σύμφωνα, εκ των οποίων ίσως αρχίζει ελληνική λέξη. Σε τέτοιες περιπτώσεις, σύμφωνα με τον κανόνα (α) θα είχαμε λανθασμένους συλλαβισμούς όπως: τά-κτ, κ.λπ. Για να αποφύγουμε τέτοια λάθη, ελέγξαμε όλες τις καταλήξεις της Νέας Έλληνικής Γλώσσας στο αντίστροφο λεξικό του Κουρμούλη [15] και προσθέσαμε τους ακόλουθους κώδικες: 6br. 6kl. 6kt. ... 6tr. 6ft..

(θ) Συνδυασμοί δίψηφων συμφώνων χωρίζονται

Πρόκειται για όρισμένους σπάνιους συνδυασμούς ξενικής προέλευσης που απο-τελούνται από δύο δίψηφα σύμφωνα όπως, π.χ., μπ-ντ (ρομπ-ντε-σάμπρ). Για την αντιμετώπιση τέτοιων περιπτώσεων, προσθέτουμε τους κώδικες: 6g2k5m2 ... 6ts5n2t. (Οι κώδικες αυτοί διαφέρουν κάπως στην περίπτωση του μονοτονικού εξ αιτίας της διαφορετικής αντιμετώπισης των ρινικών συμπλεγμάτων.)

## 4 Ένα πείραμα

Με τους βελτιωμένους κώδικες συλλαβισμού, δημιουργήσαμε μία νέα φόρμα bplain για τα αγγλικά και τα ελληνικά (ή bplain είναι μία φόρμα σαν την plain, με την διαφορά ότι επιτρέπει την συνύπαρξη πολλών γλωσσών [16]). Δημιουργήσαμε επίσης ένα αρχείο που περιέχει τις παρακάτω γραμμές:

```
\lefthyphenmin=1
\rightthyphenmin=1
\language=1

\def\mygreektext{>Eke'inh t'hn P'empth...
.
.
... to~u Ma'"iou.}
\showhyphens{\mygreektext}
\vfill\eject\end
```

(Το πλήρες κείμενο δίνεται στο Σχ. 1.) Τέλος, τρέξαμε το αρχείο με το T<sub>E</sub>X:

```
tex αρχείο &bplain hy-test
```

Από το αρχείο .log μετρήσαμε τις άστοχίες (δηλ., θέσεις στις οποίες το T<sub>E</sub>X δεν έβαλε ένωτικό) και τα σφάλματα (δηλ., θέσεις στις οποίες το T<sub>E</sub>X έβαλε κατά λάθος ένωτικό). Κατόπιν, επαναλάβαμε το ίδιο πείραμα με τους κώδικες συλλαβισμού του Χαραλάμπους [1] και του Beccari [4]. Τα αποτελέσματα, που παρουσιάζονται στον Πίνακα 1, δείχνουν σαφώς την άνωτερότητα των νέων κωδίκων. Οι νέοι κώδικες δεν παρουσίασαν κανένα σφάλμα. Οι ελάχιστες άστοχίες (3%)

πού παρουσίασαν οί νέοι κώδικες όφείλονται αποκλειστικώς και μόνο σέ όρισμένους συνδυασμούς φωνηέντων πού έκλαμβάνονται ως καταχρηστικοί δίφθογγοί, ενώ στην πραγματικότητα δέν είναι. Όπως έξηγήσαμε παραπάνω, με τούς νέους κώδικες συλλαβισμού, τó TEX έκλαμβάνει ως καταχρηστικούς δίφθογγους όλους τούς συνδυασμούς φωνηέντων πού από την γραφή είναι πρακτικώς αδύνατο νά διακρίνουμε εάν πρόκειται για καταχρηστικούς δίφθογγους ή μή.

Έκεινη την Πέμπτη, ό κύριος Ίωσήφ, ό μπαρμπέρης και μόνιμος έχθρός μας, μπήκε σά σίφουνας στην αύλή μας κι άρχισε νά ούρλιάζει ότι τó τρελλό κατσίκι μας, ή έπονομαζόμενη «Άμάλθεια», έφαγε τά νεράντζια του. Ή λάμψη στο μουστάκι του «ταχυσβαρμιστή» — ή άλλιώς «ίσοπεδωτή» — γείτονά μας είχε έξαφανιστεί και τó γόητρό του έμοιαζε νά έχει τραυματιστεί άνεπανόρθωτα. «Θά σάς κάμω μήνυσιν!», γκάριζε, «διά τó δένδρο μου τó ρημαγμένον!». Την έπομένη, γύρευε ως άποζημίωση μιá δωδεκάδα άβγά. Ό μάρμπας μας ό Άλέκος, πού είχε ανάλαβει έν μέρει τó ρόλο του πατέρα μας από τότε πού εκείνος χάθηκε στην Άλβανία, του είπε με τή χαρακτηριστική χωριάτικη προφορά του: «Δέ' μ' λές βρέ άνθρωπε, τί σ' έκαμει του κατσίκ'; Του κατσίκ' δέ ντρώει νιράτζια — έλιές τρώει!» Ό αίσχρός γείτονας επέμενε στο βιολί του και, στην κομπορυσούνη και τó θυμό του, μέρδευε λέξεις παντελώς άσχετες: «Δέν είμεθα άλλτρούσταί, κύριοι! Νά έξβγη ή οικία ύμωv εις τó σφυρί! Θά σάς μηνύσω εις τó ύποθηκοφυλάκειον, εις τó ύπουργείον! Δέν είναι δυνατόν τά προτόντα ύμωv νά γίνονται βρωσις μίας αίγός!» Είδε κι άπόειδε ή μάνα μας δανείστηκε πέντε άβγά από τόν θεió μας, έβαλε κι άλλα έφτά δικά μας και τά πήγε στο «κακοποιό μίασμα», όπως συνήθιζε νά αποκαλεί τόν κύριο Ίωσήφ. . . Την θυμάμαι ακόμα με τά μαύρα ρούχα και τó άσφράγιστο πανέρι στα χέρια της, έμπρός από τó «άνυπέρβλητον» πλατύσαλο του άρχοντοχωριάτη μπαρμπέρη — ώρα πού έδνε ό ήλιος στη θάλασσα, ένα ηλεκτροφόρο ήλιοβασίλεμα του Μαΐου.

Σχήμα 1: Τό κείμενο πού χρησιμοποιήθηκε κατά την δοκιμή τών κωδίκων συλλαβισμού.

Πίνακας 1: Συγκριτικά άποτελέσματα δοκιμής κωδίκων συλλαβισμού νέων ελληνικών.

Άρχείο κωδίκων	Σφάλματα (%)	Άστοχίες (%)
rgr-hyph.tex [1]	25	13
grhyph.tex [4]	3	16
Νέοι βελτιωμένοι κώδικες	0	3

## 5 Ποϋ θά βρεΐτε τούς νέους κώδικες συλλαβισμού

Οί νέοι κώδικες συλλαβισμού βρίσκονται σέ δύο άρχεία πού φέρουν τά όνόματα GRPhyph?.tex για τó πολυτονικό και GRMhyph?.tex, όπου ? είναι ό αριθμός τής τρέχουσας έκδοσής τους (2). Τά άρχεία αυτά μπορείτε νά τά λάβετε με ftp



ἀπὸ τὸν κατάλογο `pub/TeX/greek/hyphen` στὸν κόμβο τοῦ `*εφτ*` (`obelix.ee.duth.gr`). Θὰ τὰ βρεῖτε ἐπίσης στοὺς ἀντίστοιχους καταλόγους τοῦ CTAN.

Ἐλπίζουμε ὅτι τὰ ἀρχεῖα τῶν βετιωμένων κωδίκων συλλαβισμού γιὰ τὰ νέα ἑλληνικά θὰ ἐνσωματωθοῦν στὶς μελλοντικὲς ἐκδόσεις τῆς ἑλληνικῆς ἐπιλογῆς τῆς `babel` καὶ τοῦ Ἰωμέγα. Εὐελπιστοῦμε ἐπίσης ὅτι σύντομα θὰ μᾶς δωθεῖ ἡ εὐκαιρία νὰ παρουσιάσουμε καὶ τοὺς κώδικες συλλαβισμού γιὰ τὰ ἀρχαῖα ἑλληνικά.

## Εὐχαριστίες

Ὁ συγγραφέας θὰ ἤθελε νὰ ἐκφράσει τὶς εὐχαριστίες του πρὸς τὸν κ. Σπύρο Κωνσταντάτο, λέκτορα Κλασικῶν Σπουδῶν στὸ Πανεπιστήμιο McGill τοῦ Καναδά, γιὰ τὶς πολύτιμες συμβουλές του.

## Ἐναφορὲς

- [1] Γ. Χαραλάμπους (1990), `rgrhyph.tex`, vers. 1.1. CTAN: `tex-archive/language/greek/yannis`.
- [2] Κ. Ι. Δρυλλεράκης (1994), `grkhyphen.tex`, GREEK $\TeX$ , vers. 4.0α. `ftp://laotzu.doc.ic.ac.uk/pub/tex`.
- [3] Γ. Μοσχοβάκης (1994), `gehyphen.tex`. CTAN: `tex-archive/language/greek/greektex`.
- [4] C. Beccari (1997), `grhyph.tex` (1997). CTAN: `tex-archive/language/greek/package-babel/hyphenation`.
- [5] Μ. Α. Τριανταφυλλίδης (1975), *Μικρὴ Νεοελληνικὴ Γραμματικὴ*, β' ἔκδ. Ἄριστοτέλειον Πανεπιστήμιον Θεσσαλονίκης, Ἰνστιτούτον Νεοελληνικῶν Σπουδῶν, Ἰδρυμα Μανόλη Τριανταφυλλίδη, Θεσσαλονίκη, §21–42.
- [6] S. Levy (1988), “Using Greek fonts with  $\TeX$ .” *TUGboat* **9**, 20–24.
- [7] C. Beccari καὶ A. Syropoulos (1998), “New Greek fonts and the greek option of the babel package.” *TUGboat* **19**, 419–425.
- [8] D. E. Knuth, *The  $\TeX$ book* (1989). Addison-Wesley, Reading, Massachusetts, USA, pp. 449–451.
- [9] Y. Haralambous (1992), “Hyphenation patterns for ancient Greek and Latin.” *TUGboat* **13**, 457–469.
- [10] Μ. Χ. Οἰκονόμου (1987), *Γραμματικὴ τῆς Ἀρχαίας Ἑλληνικῆς*, ε' ἔκδ. Ὄργανισμὸς Ἐκδόσεως Διδακτικῶν Βιβλίων, Ἀθήνα, §30–34.
- [11] *Νεοελληνικὴ Γραμματικὴ*, δ' ἔκδ. (1979). Ὄργανισμὸς Ἐκδόσεως Διδακτικῶν Βιβλίων, Ἀθήνα, σελ. 11–21 καὶ 211.  
Πρόκειται γιὰ ἀναπροσαρμογὴ τῆς *Μικρῆς Νεοελληνικῆς Γραμματικῆς* τοῦ

Τριανταφυλλίδη [5]. Κατά την αναπροσαρμογή αυτή, καταργήθηκε η βαρεία (˘) και άπλοποιήθηκαν πολλοί κανόνες συλλαβισμού, τονισμού και όρθογραφίας.

- [12] *Νεοελληνική Γραμματική*, η' έκδ. (1985). 'Οργανισμός Έκδόσεως Διδακτικῶν Βιβλίων, Αθήνα, σελ. 11–25 και 211.
- [13] F. M. Liang, P. Breitenlohner and K. Berry (1996), PATGEN: PATern GENeration Program, vers. 2.3. CTAN: [tex-archive/systems/knuth/unsupported/texware/patgen.web](http://tex-archive/systems/knuth/unsupported/texware/patgen.web).  
Ένας μικρός οδηγός για την χρήση του προγράμματος PATGEN δίνεται στο αρχείο: Y. Haralambous (1993), “A small tutorial on the multilingual features of PatGen2.” CTAN: [tex-archive/info/patgen2.tutorial](http://tex-archive/info/patgen2.tutorial). Έπίσης, πληροφορίες για τὸ PATGEN θὰ βρεῖτε στὸ ἄρθρο: P. Sojka (1999), “Hyphenation on demand,” *TUGboat* **20**, 241–247.
- [14] Ν. Π. Ἀνδριώτης (1967), *Ἑτυμολογικὸ Λεξικὸ τῆς Κοινῆς Νεοελληνικῆς*, β' έκδ. Ἀριστοτέλειον Πανεπιστήμιον Θεσσαλονίκης, Ἰνστιτούτον Νεοελληνικῶν Σπουδῶν, Ἰδρυμα Μανόλη Τριανταφυλλίδη, Θεσσαλονίκη.
- [15] Γ. Ι. Κουρμούλης (1967), *Ἀντίστροφον Λεξικὸν τῆς Νέας Ἑλληνικῆς*, Ἀθήναι.
- [16] J. Braams (1998), `bplain.tex`, `babel`, vers. 3.6. CTAN: [tex-archive/macros/latex/required/babel](http://tex-archive/macros/latex/required/babel).

## Παράρτημα

Κατάλογος συμπλεγμάτων από κλιτές λέξεις ξενικής προέλευσης που περιέχουν ἄρρινα διψήφια σύμφωνα γκ, μπ και ντ. Μερικές απαραίτητες εξαιρέσεις παρουσιάζονται με ἑσωτερική ὀδόντωση.

---

### γκ

bg2k	
4l5g2k	
4n5g2k	
4r5g2k	
sg2k	
4t5g2k	
.giou5g2ko	γιου-γκοσλάβος
5g2krant.	Βόλγκο-γκραντ
>emi5g2kr	ἔμι-γκρές
>i5g2kou	ἱ-γκουάνα
>'intri5g2k >intri5g2k	ἴντρι-γκα
.kara5g2ki	καρα-γκιόζης
.ma5g2ki'o	μα-γκιόρος
.mpa5g2k	μπα-γκατέλα μπα-γκέτα
.mpi5g2k	μπι-γκόνια
.nt'a5g2kl	ντά-γκλα
.nta5g2kl	ντα-γκλαράς
.xe5g2k .x'e5g2k	ξε-γκάστρι
.saltip'a5g2k	σαλτιπά-γκος
.si5g2k	σι-γκούνα
.sp'a5g2k .spa5g2k	σπά-γκος σπα-γκοραμμένος
.fi'o5g2k .fio5g2k	φιό-γκος φιο-γκάκι
.qo'uli5g2k .qouli5g2k	χούλι-γκαν χουλι-γκανισμός

### μπ

4z5m2p	
4l5m2p	
4n5m2p	
4r5m2p	
sm2p	
>ala5m2prou	άλα-μπουρνέζικα
>'a5m2ak >a5m2p'ak	ἄ-μπακος
>a5m2pan	ἄ-μπανόζι
>a5m2p'ar >a5m2p~ar >a5m2par	ἄ-μπάρα ἄ-μπαρώνω
>a6m3p'ari >a6m3pari	ἄμ-πάρι
>a5m2pr	ἄ-μπρί
>ara5m2p	ἄρα-μπάς
>arlo'u5m2p >arlo~u5m2p >arlou5m2p	ἄρλου-μπα
.b'o5m2pir .bo5m2pir	βό-μπιρας βο-μπιριασμένος
.g'a5m2pi .ga5m2pi	γά-μπια
.ze'i5m2p	ζεῖ-μπέκιος
>i5m2pra	ἱ-μπραήμ
.kala5m2pa	καλα-μπαλίκι
.kal'i5m2pr .kali5m2pr	καλί-μπρα καλι-μπράρω
.ka5m2p'ad	κα-μπάδικος
.ka5m2par'e .ka5m2par'e .ka5m2pare	κα-μπα-ρέ κα-μπαρετζού
.ka5m2p'in .ka5m2p~in .ka5m2pin	κα-μπίνα κα-μπινές

.k'a5m2pot .ka5m2p'ot .ka5m2pot  
 .kaou5m2p  
 .kara5m2p  
     .kara6m3p'ol .kara6m3pol  
 .karna5m2p  
 .ke5m2p  
 .koli5m2pr  
 .kolo'u5m2pr .kolo~u5m2pr .kolu5m2pr  
 .k'o5m2pr .ko5m2pr  
     .ko6m3pr'ess .ko6m3press  
 .kontra5m2p  
 .koulo'u5m2pr .koulo~u5m2pr .koulu5m2pr  
 .kou5m2po'u .kou5m2po~u .kou5m2pou  
 .kwlo'u5m2pr .kwlo~u5m2pr .kwlu5m2pr  
 .li5m2pr  
 .l'o5m2p  
 .manit'o5m2p  
 5m2pa5m2p  
  
 5m2p'a5m2p  
     .mpa6m3p'aki .mpa6m3paki  
 .mpaske2t5m2p  
 .mpasi5m2pouz  
 5m2p'ehc. 5m2pehc. 5m2p'ehs. 5m2pehs.  
 5m2p'eh. 5m2peh. 5m2p'ehdec.  
 5m2p'ehdes. 5m2p'ehdwn.  
 5m2p'e5m2p 5m2pe5m2p  
  
 .mpi5m2p  
 5m2pogi  
 .mp'o5m2pir .mpo5m2pir  
 .mpo5m2p'ot .mpo5m2pot  
 .mpou5m2po'un .mpou5m2po~un .mpou5m2poun  
 .n'o5m2pel .no5m2pel  
 .nt'o5m2pr .nto5m2pr  
 .x'e5m2p .xe5m2p  
 >'o5m2po >o5m2po  
 .re5m2p  
 .r'e5m2p  
 .repo'u5m2pl .repou5m2pl  
 .ro5m2p  
 r'o5m2pa r'o5m2pe r'o5m2pw ro5m2p~w  
 ro5m2pa ro5m2pe ro5m2pw  
     tro6m3p tr'o6m3p  
 .rou5m2p  
 .sa2m5pot  
 .s'e5m2pr  
 .ska5m2pab  
 .sno5m2p  
 s'o5m2pa s'o5m2pe s'o5m2pw so5m2pa  
 so5m2pe so5m2pw so5m2p~w  
 .so'u5m2p .sou5m2p  
 .tabla5m2p  
 .ta5m2p  
  
     .ta6m3peram  
     .ta6m3po~url .ta6m3po'url

κά-μποτο κα-μποτίνος  
 καου-μπόης  
 καρα-μπίνα καρα-μπινιέρος  
 καραμ-πόλα  
 καρνα-μπίτι  
 κε-μπάπ κε-μπαπτζής κε-μπές  
 κολι-μπρί  
 κολού-μπρα  
 κό-μπρα  
 κομ-πρέσσα  
 κοντρα-μπάσο κοντρα-μπάντο  
 κουλού-μπρα  
 κου-μπούρι κου-μπούρας  
 κωλού-μπρα  
 λι-μπρέτο  
 λό-μπυ λό-μπι  
 μανιτό-μπα  
 μπα-μπάς μπα-μπάκας μπα-μπέσης  
 χαζο-μπα-μπάς  
 μπά-μπω  
 μπαμ-πάκι  
 μπασκετ-μπωλίστας  
 μπασι-μπουζούκος  
 Πετρό-μπεης κ.ἄ.ῶ.  
  
 μπέ-μπεης μπέ-μπα μπε-μπέκος  
 βουτυρο-μπε-μπές  
 μπι-μπερό μπι-μπελό  
 νερο-μπογιά λαδο-μπογιά  
 μπό-μπιρας μπο-μπιρασμένος  
 μπο-μπότα  
 μπου-μπούνας μπου-μπουνίζω  
 νό-μπελ νο-μπελίστας  
 ντό-μπρος ντο-μπροσύνη  
 ξέ-μπαρκος  
 ὄ-μπος ὀ-μποῖστας  
 ρε-μπέτης  
 ρέ-μπελος  
 ρεπού-μπλικα ρεπου-μπλικάνος  
 ρο-μπότ ρο-μποτική  
 ρό-μπα γχαρνταρό-μπα  
 παλιόρο-μπα  
 τρομ-πόνι τρόμ-πα  
 ρου-μπίνι  
 σα-μποτάζ σα-μποτάρω  
 σέ-μπρος  
 σκα-μαβία  
 σνο-μπαρία σνο-μπισμός  
 σό-μπα ζυλό-σομπα  
  
 σού-μπιτος  
 ταβλα-μπάς  
 τα-μπούρι τα-μπού τα-μπλάς τα-μπέλα  
 τα-μπακιέρα  
 ταμ-περαμέντο  
 ταμ-ποϋρλο

.te5m2pes  
 .tz'a5m2p .tza5m2p  
 .tzana5m2p  
 .tzi5m2p  
 .tzou5m2p  
 .t'o5m2pr .to5m2pr  
 .tra5m2p'al .tra5m2pal  
 .tra5m2po~uk .tra5m2po'uk .tra5m2pouk  
 .ts'a5m2pa .tsa5m2pa  
 .tsi5m2po'uk .tsi5m2pouk  
 .tsi5m2po'us .tsi5m2pous  
 .f'a5m2prik .fa5m2prik  
 .fra5m2p  
 .qa5m2p  
 .q'o5m2p  
 .qontro5m2pa

τε-μπεσίρι  
 τζά-μπα τζα-μπατζής  
 τζανα-μπέτης τζανα-μπετιά  
 Τζι-μπουτί Τζι-μπαλί  
 τζου-μπές  
 τό-μπρος το-μπροσύνη  
 τρα-μπαλά τρα-μπαλίζομαι  
 τρα-μπούκος τρα-μπουκισμός  
 τσά-μπα τσα-μπατζής  
 τσι-μπούκι  
 τσι-μπούσι  
 φά-μπρικα  
 φρα-μπαλάς  
 χα-μπαίρι χα-μπερί  
 χό-μπυ χό-μπι  
 χοντρο-μπαλάς

ντ

b5n2t  
 z5n2t  
 l5n2t  
 mn2t  
 r5n2t  
 sn2t  
 >a5n2t'atz  
 >atra5n2t'e >atra5n2t'e  
 .b'i5n2te .bi5n2te  
 .giaqou5n2t  
 .gibe5n2t  
 .gk'ai5n2t .gk'a"i5n2t .gka"i5n2t  
 .likbi5n2tar  
 .ka5n2ta"i .ka5n2ta"i  
 .ka5n2tr  
 .ma"i5n2t  
 .ma5n2t'am .ma5n2t'am .ma5n2tam  
 .ma5n2t'em .ma5n2tem  
 .me"i5n2t'a .me"i5n2ta  
 .mo5n2t'el .mo5n2tel  
 mo5n2t'ern mo5n2tern  
 .mpal'a5n2t .mpala5n2t  
 .mpa5n2tan  
 .mpe5n2t  
 .mpi5n2t  
 .mpou5n2ta  
 .mpou5n2tr  
 .nisa5n2t  
 .nta5n2t  
 .ntermpe5n2t  
 5n2t'iban  
 .ntou5n2to'uk .ntou5n2to~uk  
 .ntou5n2touk  
 .xe5n2t x'e5n2t  
 >o5n2t'a >o5n2t'a >o5n2t~a  
 .reze5n2t  
 .re5n2t'i .re5n2ti  
 .ro~u5n2t .ro'u5n2t  
 .se5n2t'ef .se5n2tef

ά-ντάτζιο  
 άτρα-ντές  
 βι-ντεο βι-ντεοχάμερα  
 γιαχου-ντής  
 γιβε-ντίζω  
 γκάι-ντα  
 λικβι-νταρισμός  
 κα-νταίρι  
 κα-ντρίλια  
 μαϊ-ντανός  
 μα-ντάμ μα-νταμίτσα  
 μα-ντέμι  
 μεϊ-ντάνι  
 μο-ντέλο μο-ντελισμός  
 μο-ντέρνο μεταμο-ντερνισμός  
 μπαλά-ντα  
 μπα-ντανάς  
 μπε-ντέλι μπε-ντένι  
 μπι-ντές  
 μπου-νταλάς  
 μπου-ντρούμι  
 νισα-ντίρι  
 ντα-ντά  
 ντερμπε-ντέρης  
 μπαουλο-ντίβανο  
 ντου-ντούκα  
 ξε-ντύνω ξέ-ντυτος  
 ό-ντάς  
 ρεζε-ντά  
 ρε-ντίκολο  
 ροϋ-ντος  
 σε-ντέφι

.se5n2to~u .se5n2to'u .se5n2tou  
.tsa5n2t'i .tsa5n2ti  
.fastfou5n2t  
.qa5n2to  
.q'a5n2tr .qa5n2tr

σε-ντοῦκι  
τσα-ντίλα τσα-ντίρι  
φαστφου-ντάδικο  
χα-ντούμης  
χά-ντρα χα-ντρῶν

Κατάλογος άκλιτων λέξεων ξενικής προέλευσης που περιέχουν άρρινα δίψηφα σύμφωνα *γκ, μπ και ντ*. Οι λέξεις αυτές περιέχονται στο σύνολο `\hyphenation{...}` του αρχείου `GRPhyph?.tex`.

---

ά-λέ-γχρο  
ά-μπα-ζούρ ά-μπα-ζούρ ά-μπα-ζούρ  
ά-ντί-ο  
ά-ουτ-σάι-ντερ ά-ουτ-σάι-ντερ  
γιά-ντες  
Ζιμ-πά-μπου-ε  
κα-σχα-ντέρ κα-σχα-ντέρ  
χο-με-ντί χο-με-ντί  
Λα-μπρα-ντόρ Λα-μπρα-ντόρ  
λί-μπι-ντο  
λου-μπέν λου-μπέν λοϋ-μπεν  
μα-ρα-μποϋ μα-ρα-μποϋ μα-ρα-μποϋ  
με-ντα-γιόν με-ντα-γιόν  
μέ-ντι-ουμ μέ-ντι-α  
νε-γκλι-ζέ νε-γκλι-ζέ  
ντε-μο-ντέ ντε-μο-ντέ  
πο-γκρόμ πο-γκρόμ  
ρα-ντάρ ρα-ντάρ  
σα-μπό σα-μπό  
σλό-γκαν  
στού-ντι-ο  
τί-γχα