RUSSIAN CONJUGATION

Computer Synthesis of Russian Verb Forms

F. H. H. KORTLANDT*

0. The program presented in this paper is a computerized version of Roman Jakobson’s article “Russian Conjugation” in Word 4 (1948), 155-67 (abbr.: J). Wherever the program and Jakobson’s rules yield different results, this is indicated below in square brackets. The language in which the program is stated is ALGOL 60.

A comprehensive comment is added to the program in order to make it accessible to both linguists without an extensive knowledge of programming languages and mathematicians who are not familiar with the Russian verb system. The input of the computer consists of the program and a list of verb stems. The generated forms are the phonemic representations of the infinitive, the preterit (masc.sg., fem.sg., neuter sg., plural), the present (1st sg., 2nd sg., 3rd sg., 1st pl., 2nd pl., 3rd pl.), and the imperative (2nd sg., 2nd pl.). [The inclusive imperative is not generated because it has no special form (J 2.122). The aspect of the verb is not taken into account; therefore, the irregularly stressed form rad’ila is not generated.

There is one major difference between the language described by Jakobson and the contemporary standard language in favour of which I have decided: the unstressed alternant of the present tense suffix á is u in Jakobson’s paper (J 2.121) but a according to the modern standard.1 For the rest I have stuck rather closely to Jakobson’s description.]

* The author studied Mathematics and Linguistics at the University of Amsterdam, where he now is a member of the Staff of the Slavic Department. His dissertation Modelling the Phoneme: New Trends in East European Phonemic Theory was recently published by Mouton.

begin  
 This program generates the infinitive and the finite inflected forms of Russian verbs.
 integer a,b,v,g,d,e,z,x,i,j,k,l,m,n,o,p,r,e,t,u,f,x,c,S,y,A,E,I,O,U,w,
 i,l,s,i,f,i,d,l,sy,l,final,finalsons;
 integer array lexeme[1:20], stem[1:20], form[1:25], des[1:2];

 boolean reflexive, drop suffix, simple desinenence, vocalic desinenence, open, jvnm, soft stem, soften,
 accented, removable accent, final syllable stressed;
 boolean procedure vowel(q); integer q;
 vowel:= q=a V q=e V q=i V q=o V q=r V q=s V q=t V q=v V q=y;

 boolean procedure soft(q); integer q;
 soft:= q=j V q=q V q=s V q=t V q=y;

 boolean procedure labial(q); integer q;
 labial:= q=p V q=b V q=f V q=v V q=m;

 procedure omit; f[i]:=f[i]-1;
 procedure change(w); integer w; form[f[i]]:=w;
 procedure add(w); integer w; begin f[i]:=f[i]+1; form[f[i]]:=w end;
 procedure insert(w); integer w; begin f[i]:=f[i]+1; form[f[i]]:=form[f[i]-1]; form[f[i]-1]:=w end;
 procedure omits; s[i]:=s[i]-1;
 procedure changes(w); integer w; stem[s[i]]:=w;
 procedure adds(w); integer w; begin s[i]:=s[i]+1; stem[s[i]]:=w end;
 procedure inserts(w); integer w; begin s[i]:=s[i]+1; stem[s[i]]:=stem[s[i]-1]; stem[s[i]-1]:=w end;

 procedure add desinenence; begin integer h; for h=1 step 1 until di do add(des[h]) end;
 procedure transfer; begin integer h; for h=1 step 1 until si do form[h]:=stem[h]; f[i]:=s[i] end;
 procedure stress(w); integer w; w:=w+2;
 procedure unstress(w); integer w; if w>36 then w:=w-27;

 procedure stress last vowel;
 begin integer h; boolean unstressed; unstressed:=true; for h:=f[i] step -1 until 1 do if unstressed AND vowel(form[h]) then unstressed:=false;
 begin stress(form[h]); unstressed:=false end;

 procedure accent;
 begin integer h; if removable accent AND vocalic desinenence then
 begin for h=1 step 1 until f[i] do if vowel(form[h]) then unstress(form[h]); stress(des[1]); add desinenence
 end else if accented then
 begin if open/syl>1 V ((final=m V final=e) V (final=m V jvnm) V (open V jvnm) V des[2]=a) AND vocalic desinenence then
 begin if simple desinenence AND vowel(des[1]) then stress(des[1]); else stress last vowel; add desinenence
 end else
 begin add desinenence; stress last vowel end;
 end else add desinenence
 end accent;
1. THE PROGRAM

The program consists of a number of declarations, which serve to define certain properties of the variables and procedures used in the program and to associate them with identifiers, and a number of statements (units of operation), which begin in line 65 (label: CONSTANT VALUES). There are several kinds of statements. An assignment statement (e.g., a := 10) serves for assigning a value to a variable or procedure identifier. A go to statement (e.g., goto INPUT) interrupts the normal sequence of operations: the next statement to be executed will be the one having the indicated label. A conditional statement (e.g., if w = 101 then EXIT else ...) causes a statement to be executed or skipped depending on the value of a boolean expression. (The value of a boolean expression is either true or false.) A for clause (e.g., for h := l step 1 until 20 do) causes the statement which it precedes to be repeatedly executed (i.e. for h = 1, 2, ..., 20). A procedure statement (e.g., phonemics) serves to invoke the execution of a procedure body. This procedure body can be found in the corresponding procedure declaration (i.e. the lines 41-55 of the program), which contains a number of statements that may be preceded by a number of declarations of local identifiers (i.e.: h, q, no vowel, unvoice).

2. COMMENT ON THE PROGRAM

The numbers refer to the respective lines of the program.

2. a, ..., S represent the phoneme inventory of Russian in the order of the Cyrillic alphabet (a, ..., s); y represents the feature of softness of the preceding consonant ('); A, ..., U represent the vowels under stress (ā, ..., ū); h and w are indices.

3. li, si, fi, di are indices that indicate the length of the relevant part of the array lexeme, stem, form, des respectively; syl is the number of syllables of the input stem; final and finalcons are the final symbol (= phoneme or the feature of softness) resp. the final consonant (or the feature of softness) of the input stem.

4. The array lexeme contains the input stem and remains intact during the execution of the program until a new input stem is read.

Changes that have bearing on a whole tense (e.g., *p'isát*', *p'išů*) are carried out in the array *stem*, while those affecting only one form (e.g., *l'ub'it'*, *l'ublů*) are carried out in the array *form*. The array *des* contains the desinence until it is added to the array *form*. The declaration of a separate array *des* is necessary in view of the accent pattern of the Russian verb, cf. below (procedure *accent*). The suffix of reflexivity (if required) is added in the output procedure. [In Jakobson's paper reflexive verbs are not taken into account.] The numbers in square brackets indicate the lower and upper bound of the subscripts.

5, 6. The booleans *reflexive* and *dropp suffix* indicate the reflexivity of the verb and the dropping of the suffix *nu* in preterit forms respectively, see line 72 below. The other booleans correspond to Jakobson's categories, cf. the assignment statements in 78-83 below; *jvnm* stands for Jakobson's 'narrowly closed', *soften* reflects the condition for softening of hard consonants. [The categories of simple and vocalic desinence affect the resulting verb forms only through the procedure *accent*. Since this procedure is not used in the generation of the infinitive (84-106 below), these boolean variables are not assigned their proper, 'Jakobsonian' logical values during the execution of this part of the program: they acquire their preterit value (*false*) in line 78 already. Of course, this is only a matter of programming economy and has nothing to do with the linguistic facts.]

7-12. The meaning of these boolean procedures is self-evident: *q* is a vowel if *q* = *a* or *q* = *e* or ..., and otherwise it is not, etc. (The symbols ∧, ∨, ¬ mean and, or, not respectively.) In this program, as in Jakobson's paper (fn. 5 on p. 157), *soft* has the meaning of 'palatal or palatalized'. [There is one slight difference: the phonemically relevant softness of non-palatal sounds is here expressed by the separate symbol *y*, so strictly speaking only this feature is *soft* in the sense of the program while the preceding consonant is not, cf. line 10.]

13. The procedure *omit* reduces the length of the relevant part of the array *form* by one, so that its effect is the omission of the last relevant symbol of this array.

14. The procedure *change* assigns to the last relevant symbol of the array *form* the value of the indicated symbol, so in effect it changes the (value of the) last relevant symbol of the array *form*. This procedure is equivalent to *omit* plus *add*.

15. The procedure *add* lengthens the relevant part of the array *form* by one and fills this place with the (value of the) indicated symbol. So in effect it adds the indicated symbol to the relevant part of the array *form*. 
16. The procedure *insert* lengthens the relevant part of the array *form* by one and *inserts* the indicated *symbol* before the last relevant *symbol* of the array *form*. It is equivalent to *omit* plus *add* (the indicated *symbol*) plus *add* (the omitted *symbol*).

17-20. These procedures are identical with the ones in 13-16 but for the array that is affected: here the array *stem* is altered, instead of the array *form*.

21. The procedure *add desinence* adds the relevant part of the array *des* to the relevant part of the array *form*.

22. The procedure *transfer* makes the array *form* equal to the array *stem* as far as the elements of the array are relevant.

23. The procedure *stress* changes *a* into *A* etc., see 65, 66 below (CONSTANT VALUES).

24. The procedure *unstress* changes *A* into *a* etc. The if clause is necessary because of the for statement in line 32: if the desinence is stressed, all vowels of the stem are unstressed. (Alternative solutions for this little problem can easily be found and some of them are quicker but they all yield more complicated boolean expressions. Since the number of vowels in the input stem hardly ever exceeds three, I have chosen here for simplicity of the program rather than for rapidity in the execution.)

25-29. This procedure has an obvious meaning. A local boolean is introduced for the sake of clarity, but it can be dispensed with if use is made of a go to statement:

```
for h := fi step —1 until 1 do if vowel(form[h]) then
begin stress(form[h]); goto LABEL
end; LABEL:
```

30. The procedure *accent* adds an accent to unstressed verb forms. Within this procedure the desinence is added to the array *form*.

31, 32. If the desinence begins with a vowel the stress of verbs having a removable accent is on this vowel (J 2.61).

33-37. For verbs with a mobile accent (unaccented input stem) the following rules hold (J 2.62):

(a) verbs with open polysyllabic input stems stress either the simple desinence, or the preceding vowel if the desinence is complex: *p'isa-*‘write’, *p'isú*, *p'íšut*, *p'íší*;

(b) if the desinence does not begin with a vowel verbs with an input stem ending in *j, v, n, m* or a vowel [= all but the BROADLY closed input stems] stress the vowel that precedes the desinence, with the limitation that in the fem.sg. preterit (suffix *la*) only verbs with nonsyllabic input stems ending in a nasal follow this rule (J 2.24): *žda-*‘wait’, *ždalá*, *ždál’i*,


procedure phonemize;
begin Integer h,q; boolean no vowel;
procedure unvoice(w); integer w;
if w=9 then w:=k else if w=9 then w:=q else if w=v then w:=f else
if w=d then w:=t else if w=v then w:=c;
no vowel:=true; for h:=1 step 1 until fi do if vowel(form[h]) then no vowel:= false; if no vowel then insert(0);
if (form[f]:=j ∧ Form[i]=C)∧Form[i]=0∧Form[i]=I) then Form[i]:=E;
for h:=1 step 1 until fi do if sort(form[h]) then
begin if Form[h]=V ∧ Form[h]=g ∧ Form[h]=w ∧ (Form[h]=x ∧ V ∧ Form[h]=i) then
end else if form[h]=o then form[h]:=a else if form[h]=y ∧ (form[h]=x ∧ V ∧ Form[h]=i) then
begin for q:=h step 1 until fi-1 do form[q]:=form[q+1]; omit
end; if form[f]=y then
begin unvoice(form[f]=i)); if form[f]=c then omit
end else unvoice(form[f]=i))
end phonemize;

procedure output;
begin Integer h; for h:=2 step 1 until fi do
begin if form[h]=36 ∧ Form[h]=33 then
begin if vowel(form[h]) then PUSYM(126) else PUSYM(127); PUSYM(form[h]=27)
end else PUSYM(form[h])
end; if reflexive then
begin if form[f]=t then PUSYM(c) else PUSYM(s); if vowel(form[f]) then PUSYM(y) else PUSYM(a)
end reflexive
end output;

CONSTANT VALUES: a:=10; b:=11; c:=12; d:=13; e:=14; f:=15; g:=16; h:=17; i:=18; j:=19; k:=20; l:=21; m:=22; n:=23; o:=24; p:=25;
r:=26; s:=27; t:=28; u:=29; v:=30; w:=31; x:=33; z:=35; A:=37; B:=41; C:=45; O:=51; U:=57; S:=55; Z:=52; y:=120;
BEGIN: li:=si:=fl:=syl:=0; for h:=1 step 1 until 20 do lexeme[h]:=0; for h:=2 step 1 until 8 do stem[h]:=form[h]=93; RUNOUT; PUNC; PUNC; PUSYM(); PUSYM(2); reflexive:=drop suffix:=false;
INPUT: w:=RESYM; if w=101 then EXIT else if w=119 then goto INPUT else PUSYM(w); if w=65 then
begin for h:=1 step 1 until li do stem[h]:=lexeme[h]; si:=li
end else
begin if w=121 then reflexive:=true else if w=93 then drop suffix:=true else if w=99 then
begin li:=li+1; if li>80 then goto BEGIN else if w=126 ∨ w=127 then
begin lexeme[li]:=syl; if li=11 then
end else lexeme[li]:=lexeme[li]+w
end; goto INPUT
end input;
final:=lexeme[li]; simple determiner:= vocalic determiner:= false; open:= vowel(final);
\}v:= final=S ∨ final=M ∨ final=M; finalcons:= \} open then lexeme[li]=1 else final;
soft stem:= sort(finalcons); accent:= final syllable stressed:= false; for h:=1 step 1 until li do if vowel(lexeme[h]) then
begin syl:=syl+1; if lexeme[h]=36 then accent:= final syllable stressed:= true else final syllable stressed:= false
end;
removable accent:= \}v ∧ final syllable stressed;
soften:= "soft stem ∧ \}syl>1 ∧ (final=a\}final=\}final=\}final=0);
etc., but žm- 'press' and žn- 'reap' both žála;

(c) otherwise the last vowel of the verb form is stressed: v'od- 'conduct', v'idú, v'idú, v'ól, v'ilá, v'ilí.

N.B.: Monosyllabic verb forms are neither stressed nor unstressed for the presence of either of these characteristics requires the presence of the other elsewhere in the word. The program, however, necessarily distinguishes between stressed and unstressed vowels, so a choice is to be made. Since the vowel oppositions that exist in monosyllabic words are the same as can be found in the stressed syllables of polysyllabic words, I have chosen to consider the vowels of monosyllabic verb forms as stressed. (The same choice was made by Jakobson.) Therefore, a 'stressed' vowel is the stressed vowel of a polysyllabic verb form or the only vowel of a monosyllabic verb form.

39. In other cases the accent is not changed.

41. The procedure phonemics contains the morphophonemic rules of the Russian verb.

43-45. The procedure unvoice is declared within the procedure phonemics and therefore has a local character. Its meaning is obvious.

46. If a verb form contains no vowel, an ķ is inserted before its final consonant (J 2.5), e.g., žg- 'burn', žók, žglá; id- 'go', šól, šlá. Instead of a local boolean (no vowel) a go to statement could have been used: if vowelform[1] then goto LABEL; insert(Ô); LABEL:

47. This ķ, like any other ķ, is replaced by é if it is preceded by a soft consonant and followed by word-final ĵ or ě: p'j- 'drink', p'ěj; p'ók- 'bake', p'ěč. [Jakobson's rule on inserted vowels runs as follows (J 2.5): "A vowel is inserted within a nonsyllabic full-stem before a nonsyllabic desinence and, if this stem ends in ĵ, before any consonantal desinence. The inserted vowel is é in the Infinitive, ķ elsewhere." According to this rule, the imperative of p'j- is *p'ój instead of p'ěj. A rule changing both this ķ and the ķ of *p'óč into é is lacking in Jakobson's description, though the necessity of having such a rule is explicitly referred to in the introduction (J 1.32): "Ô is not admitted between two soft consonants of the stem". However, this formulation of the rule is incorrect because there exist such forms as l'óža 'lying' and jóžita 'to cower'.]

48, 49. If preceded by a soft consonant, unstressed e and ō are reduced to i. The same holds true for unstressed a if it is not the final or prefinal phomeme of a word form, e.g., tr'as- 'shake', tr'ísút, but v'id'e- 'see', v'id'at. [No vowel reduction rules are explicitly stated in Jakobson's

8 See C. L. Ebeling, "On Accent in Dutch and the Phoneme /ʃ/", Lingua 21 (1968), 135f. on the essence of configurational features.
50, 51. Elsewhere (i.e. word-initially and after a hard consonant) unstressed o is reduced to a. The feature of softness y is omitted before r, where the softness is lost, and before j, where it is automatic. [Jakobson maintains the soft sign ' before j. So did I only in the input stem. In fact, it could have been omitted there as well for it does not convey any information at all. If it had actually been omitted, it would have to be inserted before the vowel insertion in the infinitive (and preterit) and the imperative (pj-/p'j- 'drink', p'it', p'ej). Since any decision on the input stem is more or less arbitrary, for once I have preferred to stick to Jakobson's notation.]

52-54. Word-final consonants are unvoiced; in addition, word-final velars lose their softness: l'og- 'lie down', imperative l'ak. [This rule is not explicitly stated in Jakobson's paper.]

56. The procedure output punches the array form into the output tape and adds the required suffix to reflexive verb forms.

57. Every verb form is preceded by three spaces, cf. the assignment statement in line 68.

58-60. Within the program capital letters (which have values between 36 and 63, see line 66 of the program) stand for stressed vowels and palatal consonants (see line 2 of the program). However, in the output tape (and also in the input tape, see below) stressed vowels are marked by underlining and the hácek on a palatal consonant is shown as a vertical line through the letter. The punching of either of these symbols precedes the punching of the letter which is marked by it (like the printing of an accent on an ordinary typewriter). So two symbols are punched instead of a capital letter: the 'accent' is punched first (symbol no. 126 which is underlining if it is a vowel, otherwise symbol no. 127 which is the vertical line) and followed by the corresponding non-capital letter, the value of which is 27 less than the value of the capital letter (cf. the procedures stress and unstress above). The procedure PUSYM punches the symbol with the indicated value into the output tape.

61, 62. The suffix of reflexivity is ca after t, sa after any other consonant, s' after a vowel: sm'éjá"- 'laugh', sm'iðóca, sm'ijálša, sm'ijálšas'.

65, 66. Here the statements of the program begin. The letters used in the program are assigned constant values. These values are equal to the values that correspond with the letters in the procedures RESYM and PUSYM, so a is symbol no. 10, etc. There is one exception: symbol no. 120 is ' instead of y, so an apostrophe in the input tape is read as y and
PUSYM (y) punches an apostrophe into the output tape. This concession to ALGOL is necessary because an apostrophe cannot be an identifier and therefore such expressions as add(') are senseless.

67, 68. This is where the execution of the program really gets started. The indices that indicate the lengths of the relevant parts of the arrays and the number of syllables of the input stem are set equal to zero. The elements of the array lexeme are assigned the value 0 because of the assignment statement in line 75. The elements of the arrays stem and form must have some value as soon as they may appear in an if clause (e.g., in 140-60). For the sake of convenience and in order to have every verb form preceded by three spaces (see line 57), all these elements are assigned the value of a space, which is 93. RUNOUT punches 20 cm of blank tape. PUNLCR punches the value of a new line and carriage return. When the input stem is punched into the output tape, it is preceded by symbol no. 74, i.e. >, and two spaces. (PUSPACE(2) is equivalent to PUSYM(93); PUSYM(93).)

69, 70. The procedure RESYM reads a symbol of the input tape and w is assigned the value of this symbol. Symbol no. 101, i.e. ], marks the end of the input. When this symbol is read the procedure EXIT makes an end to the execution of the program. The new line and carriage return symbol, which has value 119, is skipped and the execution starts once more from the label INPUT at the beginning of this line, so the next symbol of the input tape is read. Any other symbol (including a space) is punched into the output tape. Symbol no. 65, which is a hyphen, marks the end of the input stem. When this symbol is read, the array stem is made equal to the array lexeme, which then contains the letters of the input stem; otherwise, the lines 72-75 are executed and after line 76 the execution starts again from the label INPUT, that is, the next symbol of the input stem is read.

72. The reflexivity of a verb is marked by the presence of the symbol ", which has value 121, in the input stem, e.g., sm'ejaw (or sm'ejá-)'laugh'. When this symbol is read, the boolean reflexive is assigned the value true. Otherwise, the value is false because of the assignment statement in line 68. (A boolean variable must have one of these two values as soon as it appears in an if clause.) The suffix nu is in brackets if it is dropped in the preterit (cf. J 2.23), e.g., gás(nu)- 'be extinguished'. The opening bracket (= symbol no. 98) causes the boolean drop suffix to be assigned the value true, the closing bracket (= symbol no. 99) is skipped.

73-75. Any symbol which has not been mentioned thus far is now added to the array lexeme, so the index li is increased by one. If li
INFinitive: if jvnm then
  begin omits; if final=j then
    begin if syl=0 then adds(1) else if syl=1 then
      begin if stem[si]=o then begin changes(y); adds(e) end else if stem[si]=O then changes(I)
      end
    end else if (final=n ∨ final=m) ∧ syl=0 then adds(a)
  end truncate jvnm; transfer;
  if final=k ∨ final=g then
    begin change(c); if 'accented then stress last vowel
    end else
    begin if final=t ∨ final=d ∨ final=s ∨ final=b then
      begin if form[fi-1]=s then omit else change(s)
      end else if syl=0 ∧ final=r then
        begin insert(e); add(y); add(c)
        end; add(t); add(y); if 'accented then
        begin if form[fi-2]=s then add(1) else stress last vowel;
        deviating forms: if form[1]=I ∧ form[2]=e ∧ f'i=5 then
          begin forn=2 step 1 until 4 do form[fi]:=form[fi1]; omit
            begin omit; insert(s); add(y)
            end deviating forms
        end
      end; phonemics; FUNLGR; if reflexive ∧ form[fi]=y then omit; output;
    end;
  
FLister: if final=t ∨ final=d then omits else if drop suffix then si:=si-2 else if syl=0 ∧ final=r then
  begin inserts(O); accented:=removable accent:=true
  end; if si=1 then changes(S); transfer;
  if vowel(form[fi]) ∨ si=1 then adds(1); if 'accented then stress last vowel; phonemics; PUSPACE(3); output;
    begin lexeme[li]:=stem[li]; k; li:=si:=h
    end tolko; transfer;
  begin form[li]=s; des[2]:=A
  end pryad; if final=j ∧ ((form[1]=b ∨ form[1]=s) ∧ syl=0 ∨ form[fi]=e ∧ syl=1) then
    begin stress last vowel; add desinence
    end else accent; phonemics; output;
    transfer; des[2]:=ο; accent; phonemics; output;
  if form[fi]=0 then begin change(y); add(1) end else begin change(y); add(1) end; output;
becomes too large (cf. the declaration of the array \textit{lexeme} in line 4), the execution starts again from the label \texttt{BEGIN} as if line 200 had been reached. Otherwise, \textit{lexeme}[li] is assigned the value of the symbol that has been read. There is one complication: stressed vowels and palatal consonants are compressed into one symbol, which is the corresponding capital letter. This means that when the stress mark of a vowel (symbol no. 126) or the háček of a consonant (symbol no. 127) is read, the symbol to be read next is to be changed into a capital letter, i.e. to be increased by 27 (cf. the comment on 58-60). This compression of stressed vowels and palatal consonants is not really necessary, but it induces considerable simplifications in a number of statements, e.g. in the procedures \textit{stress} and \textit{unstress}, the softening rules of the present tense, and the condition for the desinence of the imperative (line 197). The palatalized consonants could similarly be compressed, but here the need to do so is less urgent and it would necessitate a number of changes in various statements that would not further the readability of the program.

76. Now the next symbol of the input stem is to be read unless the previous symbol was a hyphen.

78-82. The meaning of these assignment statements will be obvious to anyone who has succeeded in reading the program so far. The values of the booleans \textit{accented} and \textit{final syllable stressed} are set equal to \texttt{false} before the for statement that counts the syllables (the number of syllables is equal to the number of vowels) and in which the values of these booleans are set equal to \texttt{true} when a stressed vowel occurs. Then \textit{final syllable stressed} becomes \texttt{false} again if there follows an unstressed vowel. Open and broadly closed input stems have a removable accent if their final syllable is stressed (J 2.61).

83. The final consonant of hard polysyllabic input stems ending in \textit{a} or \textit{o} is softened in the present and the imperative (J 2.42.A), see 161-69.

84-90. These lines contain the changes of the stem that are common to the infinitive and the preterit forms. Input stems in \textit{j}, \textit{v}, \textit{n}, \textit{m} drop their terminal phoneme (J 2.221). Nonsyllabic stems in \textit{j} replace this \textit{j} by \textit{i}, e.g., \textit{p'j}- 'drink', \textit{p'it'}, while monosyllabic stems replace the vowel \textit{o} before the dropped \textit{j} by \textit{i} [if stressed, otherwise by \textit{e} which (unlike \textit{i}) palatalizes the preceding consonant], e.g., \textit{mój}- 'wash', \textit{mit'} (cf. \textit{mójit}) [but \textit{pój}- 'sing', \textit{p'et'} (cf. pajót)]; nonsyllabic stems in \textit{n}, \textit{m} replace this nasal by \textit{a}, so \textit{žm}- 'press' and \textit{žn}- 'reap', both \textit{žát'} (J 2.24). After these changes the stem is transferred to the array \textit{form}, where the infinitive is generated. The array \textit{stem} is preserved for the formation of the preterit in 107 ff.
91, 92. Verbs with an input stem ending in a velar have an infinitive ending in ĝ (J 2.112 and 2.222), e.g., p'ok- 'bake', p'ēĉ.

94, 95. All terminal dentals and labials of the broadly closed input stems coalesce into s before the infinitive desinence (J 2.3), e.g., gr'ob- 'row', gr'ist'i. [This s coalesces with a preceding s into one phoneme: rost- 'grow', rast'i.]

96, 97. The infinitive of nonsyllabic stems in r shows polnoglasie (J 3.1), e.g., t'r- 'rub', t'ir'ēt'.

98, 99. The infinitive desinence is t'i in verbs with unaccented input stems if the desinence is preceded by a consonant [which must of course be s], t' elsewhere (J 2.112). The desinence is here directly added to the array form.

100-104. There are two exceptions (J 3.1): id- 'go', it'i instead of *ist'i and kl'an- 'curse', kl'âst' instead of *kl'ât'.

106. The infinitive is punched after a new line and carriage return symbol. If it ends in t' and the verb is reflexive, the soft sign is omitted before the suffix ca (which is added in the output procedure, 61 f.), e.g., sm'eji'- 'laugh', sm'ijâteca.

107, 108. Now we return to the array stem. If the input stem ends in t or d this dental stop is dropped before the preterit desinences (J 2.222). So is the suffix mu if it was in brackets in the input stem (J 2.23), cf. line 72. Nonsyllabic input stems in r insert a stressed o in the preterit forms (J 2.5), e.g., t'r- 'rub', t'or, t'orla. (The accent is removable, however, because the desinence is stressed in the present and the imperative where there is no inserted vowel.)

109. The stem id- 'go', which lost its dental stop in line 107 so that its length has been reduced to one, is replaced by s in the preterit: sôl (vowel insertion in line 46 of the procedure phonemics which is invoked in line 110), šlâ. After these changes, which are common to all preterit forms, the stem is transferred to the array form where the desinence and the accent are added (if necessary).

110. The masc.sg. desinence is l after a vowel and zero after a consonant (J 2.111). [The form sôl is not mentioned by Jakobson.] The output is preceded by three extra spaces.

111-13. [The correct preterit forms of the verb tolok-/tolk- 'pound', talôč, talôk, talklā, tâłôt are not found in Jakobson's paper. The forms of the infinitive and the masc.sg. preterit are derived from one stem and all other forms from the other. The alternation cannot be automatic because there exist such verbs as volok- 'drag', valôč, valôk, valaklā, valâčôt and môlk(nu)- 'become silent', môlknut', môlk, môlkla, môlknit.
I have not been able to find a more elegant solution to this problem.]

Once more the stem is transferred to the array form.

114-18. The fem.sg. desinence is *la (J 2.111). The form pr’ila, derived from the input stem pr’ad- ‘spin’, is irregularly stressed. [According to Jakobson the input stem is pr’ad- and the irregularly stressed forms are pr’dla, pr’dli (J 3.2). But this is incorrect because his input stem would generate the infinitive *pr’ist’i instead of pr’ast’, a fact which he clearly failed to realize.] The forms b’ila and šila, derived from the input stems b’j- ‘beat’ and šj- ‘sew’ respectively, are also irregularly stressed [these verbs are not even mentioned by Jakobson], and so is the form p’ela from poj- ‘sing’. [The irregular stress pattern of the latter verb is not stated explicitly though it can be derived from the form that is mentioned in J 3.2 because of its vowel alternation.]

119. The neuter sg. desinence is lo (J 2.111). It is necessary to invoke
the procedures transfer, accent, phonemics once more because the stress is often different from the stress on the previous form and the proper quality of the unstressed vowels is lost in the procedure phonemics.

120. The plural desinence is l’i (J 2.111). Since this form is for the rest identical with the previous one, it is unnecessary to invoke the procedures of line 119 again.

121. For the generation of the present and the imperative forms the original input stem is once more transferred to the array stem.

122-27. The present of the verbs daj- ‘give’ and jéd- ‘eat’ is so irregular (J 3.3) that a PUTEXT seems to be justified. Any other solution would require a great number of ad hoc rules. The procedure PUTEXT punches the indicated string into the output tape.

128. The l sg. desinence of all other verbs is u (J 2.121). Open input stems lose their final phoneme before the present and imperative desi-
nences (J 2.21).

129-35. The present stem of verbs with an input stem in ova ends in uj (J 2.24). If the final syllable of the input stem is stressed, the stress moves to the desinence in two-syllabic stems only, otherwise to the preceding vowel, e.g., ková- ‘forge’, kujú, but darová- ‘grant’, darúju. There is one exception (J 3.2): dn’ová- ‘spend the day’, dn’úju. If the final a of the input stem is not stressed, the accent remains unchanged.

136, 137. The present stem of verbs with an input stem in aváj ends in aj and the stress moves to the desinence (J 2.23): daváj- ‘give’, dajú. The input stem reappears in the imperative daváj, cf. below (line 194).

139. The following verbs show an irregular alternation (J 3.2):4

4 Professor F. E. J. Kruseman Aretz pointed out to me that the execution of the
Discrepancy between the prevocalic and the preconsonantal stem shape:

if stem[1]=\text{b} then begin changes(g); finalcons:=g end else

if stem[1]=\text{v} then begin adds(y); finalcons:=y; soft stem:=true end else

if stem[1]=\text{\=a} then begin inserts(s); end else

if stem[1]=\text{\=s} then begin changes(d); soft stem:=false end else

if stem[1]=\text{\=t} then begin changes(e); adds(e); end else

if stem[1]=\text{\=u} then begins(t); end else

if stem[1]=\text{\=d} then begin inserts(o); end else

if stem[1]=\text{\=\=a} then removable accent:=false else

multiple discrepancy:

if stem[1]=\text{\=e} then begin inserts(o); adds(y); syl:=2; soft stem:=true end else

if stem[1]=\text{\=i} then begins(t); end else

if stem[1]=\text{\=s} then begins(t); end else

if stem[1]=\text{\=d} then begins(t); end else

if stem[1]=\text{\=\=a} then removable accent:=false end else

if stem[1]=\text{\=\=e} then begins(t); end else

if stem[1]=\text{\=s} then begins(t); end else

if stem[1]=\text{\=d} then begins(t); end else

if stem[1]=\text{\=\=a} then removable accent:=false end else
RUSSIAN CONJUGATION 65

140. b’ezá- ‘run’, b’igú;
143. spa- ‘sleep’, spl’ú (palatalization in line 177);
144. r’ov’é- ‘roar’, r’ivú;
145. sm’ejá”- ‘laugh’, sm’ijús”; ržá- ‘neigh’, ržú;
146. 147. ropta- ‘grumble’, rapššú, etc. (palatalization in line 163);
148. slá- ‘send’, šl’ú;
149. jéxa- ‘drive, go’, jédu;
150. molo- ‘grind’, m’il’ú (palatalization in line 167);
151. br’ij- ‘shave’, br’éju [here, as in all other verbs which show discrepancy, I have chosen the infinitive stem shape as the representation of the input stem; the vowel alternation of the verb poj- ‘sing’ is generated by the program];
152. zva- ‘call’, zavú;
153. bra- ‘take’, b’irú; dra- ‘tear’, d’irú;
154. l’éz- ‘climb’, l’ézu; here the verb mog- ‘can’ must be added, which has the accent pattern (but not the palatalization) of open polysyllabic input stems [this verb is not mentioned by Jakobson];
156. gna- ‘drive’, gan’ú;
157. stlá- ‘spread’, st’il’ú [the accent of Jakobson’s stlalá, which requires an unstressed input stem, is confirmed by none of my informants and rejected by Avanesov and Ožegov in their Russkoe literaturnoe proiznošenie i udarenie (Moscow, 1960), p. 567];
158. l’og- ‘lie down’, l’ágú;
159. s’éd- ‘sit down’, s’ádu;
160. bi- ‘be’, búdú (the i of the input stem has already been omitted in line 129).
161. If the condition of line 83 is fulfilled, the final consonant of the stem is softened (J 2.42.A). Examples:
162, 163. pláka- ‘weep’, pláču; skaka- ‘jump’, skačú; iska- ‘look for’, iššú; pr’dra- ‘hide’, pr’áču; [I prefer to write šš instead of šč in accordance with the modern standard pronunciation 6]

program takes considerably less time if special variables are introduced to be used in the discrepancy rules instead of the subscripted variables of the array stem (e.g., sl := stem [1], etc.) because the looking up of a subscripted variable takes relatively much time.

if soften \& stem[si]=j then softening:
begin if stem[si]=k \& stem[si-1]=t \& stem[si]=t then
    begin if stem[si-1]=s then begin omit; changes(S); adds(\&) end else
        changes(C)
    end else if stem[si]=g \& stem[si]=d \& stem[si]=z then
        begin omit; if stem[si]=z then changes(Z); adds(Z)
    end else if stem[si]=x \& stem[si]=s then changes(S) else
        begin if labial(stem[si]) then adds(I); adds(y)
    end
end softening; transfer;

if soft stem \& finalcons=s then substitutive softening:
begin if form[fi]=y then omit;
    if form[fi]=t then
        begin if form[fi-1]=s then begin omit; change(S); add(S) end else
            change(C)
        end else if form[fi]=d \& form[fi]=z then
            begin omit; if form[fi]=x then change(Z); add(Z)
        end else if form[fi]=s then change(S) else
            begin if labial(form[fi]) then add(I); add(y)
        end
end substitutive softening;

if des[1]>36 then add desinence else accent; phonemics; PUNCT; output;

simple desinence:= false; di:=2; des[1]:= if open \& soft stem then i else o; des[2]:=S;
if final=A\& stem[si]=\& stem[si-1]=m \& li-si> then stress(des[1]); transfer;
if form[fi]=k \& \& form[fi-1]=t then
    begin if form[fi-1]=s then begin omit; change(S); add(S) end else
        change(C)
    end else if form[fi]=g \& form[fi]=z then
        begin omit; if form[fi]=x then change(Z); add(Z) end else
    if issoft(Form[fi]) then add(y);
if des[1]>36 then add desinence else accent; phonemics; PUNCT;
if form[fi]=x \& form[2]=x \& form[3]=x then PUTEXT{ k xoti\& xoti\&} else begin output; change(t); output end sing;
if PLURAL: change(m); output; change(t); add(y); add(I); output;

if des[1]= if open \& soft stem \& finalcons=s then a else u; des[2]=t;
if final=A\& stem[si]=\& stem[si-1]=m \& li-si> then stress(des[1]); transfer;
if des[1]>36 then add desinence else accent; phonemics; output;

IMPERATIVE: simple desinence:= true; if final=A\& stem[si]=\& stem[si-1]=u then stress(stem[si-1]);
if final=x \& li-si> then begin for h:=1 step 1 until li do form[h]=lexeme[N]; fi:=li end else transfer;
if soften\& finalcons=a\& removable accent then begin omit; change(y) end else
if issoft(form[fi]) then add(y);
if (accented\& removable accent\& vowel(form[fi-1])\& vowel(form[fi-2])\& form[fi]=y) \& finalcons=s \& final=I then
    begin di:=1; des[1]=i; accent
end else if (\& accente\& final=A) \& finalcons=s then stress last vowel; phonemics; PUNCT; output;
add(t); add(y); add(I); output; goto BEGIN

end
164, 165. brizga- ‘sprinkle’, brížžu; gloda- ‘gnaw’, glažú; máza- ‘smear’, mážu;
166. paxa- ‘plow’, paxú; p’isa- ‘write’, p’išú;
169. The stem is transferred to the array form, where the changes that are peculiar to the 1 sg. are carried out.
170. If the final consonant of the stem is soft, it most often changes into a palatal consonant or a cluster (J 2.41). This softening is ‘substitutive’ in the sense that one soft consonant is replaced by another. Examples:
172, 173. m’ét’i- ‘mark’, m’écu; mst’i- ‘revenge oneself’, mššú;
174, 175. s’id’é- ‘sit’, s’ižú; jézd’i- ‘drive’, jéžžu; grož’i- ‘menace’, gražú;
176. v’is’é- ‘hang’, v’išú;
177. šumé- ‘make a noise’, šuml’ú.
180. If the desinence has acquired the stress in line 132 or 137 already, the procedure accent need not be invoked. A new line and carriage return symbol precedes the 1 sg. in the output tape.
181. The 2 sg. desinence is iš or oš (J 2.121).
182. Verbs with an input stem ending in ová or aváj stress the desinence under the same conditions as in the 1 sg., see line 132 and 137 respectively. The accent of verbs like darova- ‘grant’, darúju is not affected because the suffix u of such verbs has received the stress in line 133 and the condition of the present if clause contains an unstressed u. The stem is transferred to the array form, where stem-final velars are replaced by palatals and other stem-final consonants undergo ‘bare’ softening (J 2.42.B):
183, 184. p’ok- ‘bake’, p’ikú, p’ičós;
185. Iga- ‘lie’, Igu, Izoš; zg- ‘burn’, zgú, zżós;
188. The forms xočiš, xočit instead of *xat’iš, *xat’ît (input stem xot’é- ‘want’) are irregular (J 3.1). The 3 sg. form is identical with the 2 sg. but for its final phoneme, which is t instead of š (J 2.121).
189. The 1 pl. and 2 pl. forms end in m and t’i respectively (J 2.121); they are for the rest identical with the 2 sg. and 3 sg.
190. The 3 pl. desinence is at if the 2 sg. ends in iš and ut if the 2 sg. ends in oš [cf. J 2.121] but for the verb b’ežá- ‘run’, which has b’igút. (The combination soft stem and finalscons = g does not occur in any other verb, cf. line 140.)
191. Cf. line 182.
193. The imperative of verbs with an input stem in ová ends in új.
194. The imperative of verbs with an input stem in aváj is identical with the input stem itself, which has been preserved in the array lexeme. Other verbs derive the imperative from the present stem.
195. The verbs sipa-'scatter' and krápa-'trickle' lose the epenthetic l in the imperative (J 3.1): sip’, kráp’.
197, 198. The imperative desinence is i if the verb has a mobile accent (unstressed input stem) or a removable accent, and also if the present stem ends in two consonants; otherwise it is zero. There is one limitation: if the final consonant of the input stem is j, the imperative desinence is i only if the input stem ends in ji (J 2.122), e.g., taji- ‘conceal’, taji, but stojá- ‘stand’, stój, and p’j- ‘drink’, p’éj.
199. If the final consonant of the input stem is j and the imperative desinence is zero, the verb form still has to receive an accent. The output is preceded by a new line and carriage return symbol.
200. The plural form is obtained by adding the suffix t’i to the generated verb form. When this form has been punched into the output tape, the next input stem is to be read.

3. THE OUTPUT

The program was executed with a list of 150 verb stems on the X8 of the Mathematical Centre, Amsterdam. The results are shown below.

Amsterdam
RUSSIAN CONJUGATION

> daj-
  dat' dal dala dala dal'i
dam dat' dat' dadi'm dadi't'i dadiut
daj daj't'i

> dv'in-
dv'inut' dv'inula dv'inula dv'inula dv'inula dv'inula
dv'inu dv'in'i dv'in'i dv'in'i dv'in'i dv'inut
dv'in' dv'in't'i

> d'elaj-
d'elat' d'elal d'elal d'elal d'elal d'elal d'elal d'elal d'elajit d'elajim d'elajit'i d'elajut
d'elaj d'elajt'i

> d'en-
d'enut' d'ela d'ela d'ela d'ela d'ela d'ela d'ela d'ena'i d'eni'im d'eni'it'i d'eniut
d'en' d'en't'i

> dn'ova-
dn'ovat' dn'ova d'elava d'elava d'elava d'elava d'elava d'elava d'elava
dn'ova d'ovat' d'ova d'ova d'ova d'ova d'ova d'ova d'ova d'ova

> d'im-
d'imat' d'imala d'imala d'imala d'imala d'imala d'imala d'imala d'imala d'imala

> d'el-
d'elat' d'ela d'ela d'ela d'ela d'ela d'ela d'ela d'ela d'elut
d'el d'elat'

> jexa-
jexat' jexal jexal jexal jexal jexal jexal jexal jexal jexal

> jex-
jexat' jexal jexal jexal jexal jexal jexal jexal jexal jexal

> jexa-
jexat' jexal jexal jexal jexal jexal jexal jexal jexal jexal
RUSSIAN CONJUGATION 73

> kolo-
    kelot' kalol kalola kalola kalol'ı
    kal'ı kal'ıt kol'ım kol'ıt'ı kol'ıt kol'ut
    kal'ı kal'ıt'ı

> krap-
    krapal krapala krapala krapal'ı
    krapal'ı krapal'ı krapal'ı krapal'ı krapal'ı krapal'ı
    krapal'ı krapal'ı

> krad-
    krad'ı kral krala kral'ı
    krad'ı krad'ı krad'ı krad'ı krad'ı krad'ı
    krad'ı krad'ı

> kr'ik-
    kr'ikan'ı kr'ika kr'ika kr'ikan'ı kr'ikan'ı kr'ikan'ı
    kr'ikan kr'ikan kr'ikan kr'ikan kr'ikan kr'ikan
    kr'ikan kr'ikan'ı

> kr'ie-
    kr'iet' kr'ieala kr'ieala kr'ieala kr'ieala kr'ieala
    kr'ieala kr'ieala kr'ieala kr'ieala kr'ieala
    kr'ieala kr'ieala

> krj-
    krj krj krj krj krj krj
    krj krj krj krj krj krj
    krj krj

> kup-
    kup'ı kup'ı kup'ı kup'ı kup'ı kup'ı
    kup'ı kup'ı kup'ı kup'ı kup'ı kup'ı
    kup'ı kup'ı

> lga-
    lgat' lgal lgala lgala lgala lgala
    lgal lgal'ı

> l'ot-
    l'ot' l'ota l'ota l'ota l'ota l'ota
    l'ota l'ota l'ota l'ota l'ota l'ota
    l'ota l'ota

> l'es-
    l'es l'es l'es l'es l'es l'es
    l'es l'es l'es l'es l'es l'es
    l'es l'es

> l'o-
    l'o l'o l'o l'o l'o l'o
    l'o l'o l'o l'o l'o l'o
    l'o l'o

> l'it-
    l'it' l'ita l'ita l'ita l'ita l'ita
    l'ita l'ita l'ita l'ita l'ita l'ita
    l'ita l'ita

> l'i-
    l'i l'i l'i l'i l'i l'i
    l'i l'i l'i l'i l'i l'i
    l'i l'i

> l'i-
    l'i l'i l'i l'i l'i l'i
    l'i l'i l'i l'i l'i l'i
    l'i l'i
RUSSIAN CONJUGATION

> orat' orala orala aral'ar
> ar' or'i or'it or' im or'it' or'ut
> ar'i ar'it'

> orat' orala orala aral'ar
> ar' or'ot ar'om ar'ot' arut
> ar'i ar'it'

> pas- paslo paslo pas'l'
> pasu pas'ot pas'om pas'ot' pasut
> pas'i pas'it'

> pas- pasl pasl pas'ol pas'l'
> pasu pas'ot pas'om pas'ot' pasut
> pas'i pas'it'

> p'et' p'ela p'ela p' el'
> pa ju pajot pajom pajot' pajut
> poj poj't'

> p'ok p'okla p'iklo p'ikl'
> p'iku p'ikot p'ikom p'ikot' p'ikut
> p'ik' p'ik'it'

> p'ela p'ela p'ala p'al'
> p'ii u p'iiot p'iiom p'iiot' p'iiut
> p'ii' p'iiit'

> plakat' plakala plakala plakal'
> pla ju pla' pla't pla't' pla't'

> pl'ovu- pl'avu pl'aval pl'aval pl'aval'
> pl'uju pl'ujo pl' ujot pl' ujom pl'ujo t pl' ujut
> pl'uj pl'ujit'

> pliv- plil plila plila plil'
> pliv pliv' pliv'ot pliv'om pliv'ot' plivut
> pliv' pliv'it'

> poj- pajil pajila pajila pajil'
> paju pajim pajit pajit' pajut
> paj' pajit'
RUSSIAN CONJUGATION

> томи-
tanum tanul tanula tanula tanul'i
tanu ton'iš ton'it ton'im ton'it'i tonut
	тан'ит'i

> торжествов-
tар'иствавал тар'иствавала тар'иствавала тар'иствавал'i
тар'иствуй тар'иствуйt тар'иствуйt'im тар'иствуйt'i тар'иствуйту
тар'иствуйt'i

> тр'епи-
trip'itala trip'itala trip'itala trip'italal'i
trip'italu trip'es'it trip'es'sim trip'es'it'i trip'es'sut
trip'es'it'i

> трои-
tronul tronula tronula tronula tronul'i
tronu tron'iš tron'it tron'im tron'it'i tronut
tron'it'i

> тр'ас-
tr'as tr'as'la tr'as'lo tr'as'l'i
tr'asu tr'as'oš tr'as'ot tr'as'ot'i tr'asut
tr'as'it'i

> тр'ас'и-
tr'as'sa tr'as'sas' tr'as'sosa tr'as'sotca tr'as'soma tr'as'sotis' tr'as'utca
tr'as's's 'tr'as's'itis'

> учи-
uchu uči' učila učila učila učila
uchu uči' učim učit'i ūčat
uchu učit'i

> учитца
uchilsa učilas' učilas' učila' učilas'
učils' učilsa učitca učima učit'is' učatca
učils' učit's'

> ходи-
hox'd'it hox'd'it' hox'd'ila hox'd'ila hox'd'il'i
hox'd'itu hox'd'it' hox'd'im hox'd'it'i hox'd'it
hox'd'it'i

> хоти-
xat'el xat'el xat'el xat'el xat'el'i
xat'elu xat'el' xat'el' xat'el' xat'el' xat'el'
xat'el' xat'el' xat'el' xat'el' xat'el' xat'el'
xat'el' xat'el' xat'el' xat'el' xat'el' xat'el'
xat'el' xat'el' xat'el' xat'el' xat'el' xat'el'

> ховоти-
xaxat't' xaxatals xaxat'als xaxat'al'i
xaxatu xaxoš't xaxoš't xaxoš'tim xaxoš't'i xaxoš'tut
xaxoš't' xaxoš'tut

> храни-
xrani' xran'iš xran'ila xran'ila xran'ila xran'ila xran'ila
xran'u xran'iš xran'it xran'im xran'it'i xran'at
xran'it' xran'at

> карапун-
karapnut karapnul karapnula karapnula karapnul'i
karapnu karapn'iš karapn'it karapn'im karapn'it'i karapnut
karapn'i karapn'it'i
> +cej(nu)- 
+cejnut' +bes +cejla +cejla +cezl'1 
+cejnu +cezn'1+ +cezn'it +cezn'im +cezn'it'1 +ceznut 
+cezn'i +cezn'it'i

> citalit-
 citait' citala citala citali' 
 citaju citajl' citajlit citajim citajit'1 citajut 
 citaj' citajit'1

> st-
 st' st'1 st'ila st'ila st'1'1 
 st'Ju st'Jos' st'lot' st'jom' st'lot'i st'jut 
 st'el' st'el'1

> sum'et-
 sum'et' sum'ela sum'ela sum'el'1 
 sum'elu sum'el' st'um'lit st'um'im st'um'it'1 st'um'at 
 st'um'i st'um'it'1

>