1

I.1 INTRODUCTION

This dissertation is about a certain class of formulas of monadic second-order logic with a single binary predicate constant, the modal formulas. These formulas are of the form

$$(\forall P_1)...(\forall P_n)\phi(P_1,..., P_n, R),$$

where $P_1,..., P_n$ are unary predicate variables and $R$ is the binary predicate constant. $\phi(P_1,..., P_n, R)$ is a formula of monadic first-order logic based on $P_1,..., P_n$ with restricted quantifiers. This can be stated more formally as follows. $\phi(P_1,..., P_n, R)$ belongs to the smallest class $E$ of expressions satisfying the following four conditions,

(i) for each individual variable $x$, $P_1x, ..., P_nx$ are expressions in $E$

(ii) if $\alpha$ is an expression in $E$, then so is $\neg\alpha$

(iii) if $\alpha$ and $\beta$ are expressions in $E$, then so is $(\alpha \rightarrow \beta)$

(iv) if $\alpha$ is an expression in $E$, then so is $(\forall y)(Rxy \rightarrow \alpha)$, for any two distinct variables $x$ and $y$.

Finally $\phi$ is required to have exactly one free individual variable.

If $\alpha$ is a modal formula, we write $\overline{\alpha}$ for the universal closure of $\alpha$ taken with respect to its one free individual variable.

The exact connection between this definition of modal formulas and more traditional ones will become clear at the end of this introduction and in chapter I.2.
Modal formulas derive their interest from two sources. In the first place, according to a theorem by S.K. Thomason (cf. [23]) there exists an effective translation \( \tau \) from sentences in the language of monadic second-order logic with one binary predicate constant \( R \) to modal formulas, and a modal formula \( \delta \) such that, for all sentences \( \phi \) and sets of sentences \( \Gamma \) in this language,

\[
\Gamma \models \phi \iff \{ \tau(\gamma) \mid \gamma \in \Gamma \} \cup \{ \delta \} \models \tau(\phi).
\]

(Here \( \models \) denotes logical consequence. In the limiting case where \( \Gamma \) is empty and \( \phi \) is universally valid, we write \( \models \phi \).) H.C. Doets showed recently that an effective translation \( \delta \) exists from second-order sentences to sentences of the form \((\forall R)(\exists P)\psi(R, P)\), where \( \psi(R, P) \) is a first-order sentence in the binary predicate variable \( R \) and the unary predicate variable \( P \), such that for all second-order sentences \( \phi \),

\[
\models \phi \iff \models \delta(\phi).
\]

Combining these results it appears that the modal formulas are, in a sense, a reduction class for second-order logic. An effective translation \( \tau \) exists from second-order sentences to modal formulas such that, for all second-order sentences \( \phi \),

\[
\models \phi \iff \overline{\delta} \models \tau(\phi).
\]

The \( \overline{\delta} \) cannot be omitted here, for the set of universally valid modal formulas is recursive, whereas the set of universally valid second-order sentences is not.

The second source of interest in modal formulas lies in the well-known possible worlds semantics for modal logic. The clauses of S. Kripke's truth definition (cf. [12]) are reflected in our syntactic clauses (i),..., (iv).

From both these points of view the following question seems a
natural one. Which modal formulas are first-order definable? More precisely, fixing $L_0$ to be the first-order language with equality containing the binary $R$ mentioned above as its only predicate constant, we ask which modal formulas are logically equivalent to $L_0$-formulas. Taking this relation of logical equivalence between modal formulas and $L_0$-formulas as our object of study we are led to an obvious converse of our first question. Which $L_0$-formulas are modally definable? More precise formulations of these questions will be found in chapters I.2 and I.6.

The above questions are treated in part I which is intended to give a survey of this area of research. Part II consists of three published contributions of our own to the subject. In addition to these we mention Van Benthem [1]. Also all results in part I that are not explicitly attributed to a particular person or the folk literature are new as far as we know.

We now give a short description of part I. In the remainder of this introduction it will be shown how modal formulas as defined here are related to modal formulas defined in a more traditional (and in fact the usual) way. Moreover, a semantic characterization is given of those formulas of monadic first-order logic that have restricted quantifiers.

I.2 contains some standard notions and results to give a first impression of modal formulas. Our question about first-order definable modal formulas is stated in a precise manner. This leads to two different versions, one for modal formulas $\phi$ ("local" correspondence) and one for modal sentences $\overline{\phi}$ ("global" correspondence). Defining $M_1$ as $\{ \phi | \phi$ is a modal formula logically equivalent to some $L_0$-formula with the same free variable as $\phi \}$ and $\overline{M}_1$ as $\{ \phi | \phi$ is a modal formula for which $\overline{\phi}$ is
logically equivalent to some $L_0$-sentence) we obtain a surprising result: $M_1 \subseteq M_1$, but $M_1 \neq M_1$. (For $M_1$, cf. Segerberg [18], Thomason [24], and Sahlqvist [16]; for $M_1$, cf. Van Bentham [1].)

I.3 gives an algebraic characterization of $M_1$. In Goldblatt [7] it is shown that a modal formula is in $M_1$ iff it is preserved under ultraproducts. This is an instance of the general result that a $\mathbb{T}^1$-sentence is first-order definable iff it is preserved under ultraproducts. Goldblatt's result is sharpened here to preservation under ultrapowers. It is also proved that a set of modal formulas defines either an $L_0$-elementary class of models, or an $L_0-\Delta$-elementary class that is not $L_0$-elementary, or a class that is not $L_0-\Sigma\Delta$-elementary. Examples of all three kinds are given.

More syntactic information on first-order definable modal formulas is provided by two methods introduced in I.4. It appears that, whereas $M_1$ was the most natural class to characterize algebraically, $M_1$ is a more suitable object for study now. The first method yields "positive" results, showing certain formulas to be in $M_1$. It proceeds roughly as follows.

Call the $L_0$-formula $\psi$ a substitution-instance of the modal formula $(\forall P_1)...(\forall P_n)\phi$, if $\psi$ is obtained from $\phi$ by substituting $L_0$-formulas for the predicate variables. (But see chapter I.4 for the exact formulations.)

Clearly, a modal formula implies each of its substitution-instances. $M_1^{\text{sub}}$ is the class of modal formulas which are implied by a conjunction of their substitution-instances. It is shown that $M_1 \subseteq M_1$ and that $M_1^{\text{sub}}$ is recursively enumerable. But $M_1^{\text{sub}} \neq M_1$, as appears from an example in I.2.

Still, this method leads to a generalization of a theorem by H. Sahlqvist (cf. [16]), which was the most comprehensive result until now.

The second method yields "negative" results, showing certain
formulas to be outside of $M_1$. Here the Löwenheim-Skolem theorem is used as follows. We show that the modal formula under consideration holds in an uncountable model, for some assignment to its free variable, but that it does not hold in any countable elementary submodel of a suitable kind. A number of examples obtained in this way show that the generalisation of Sahlqvist's result referred to above is "almost" the best possible result.

A combination of the two methods leads to a complete syntactic classification of the first-order definable modal reduction principles. We do not define this notion here, but the definition is in chapter I.4. (Cf. II.2 and Fitch [6].) Many of the better-known axioms used in modal logic are modal reduction principles.

I.5 deals with particular cases where $R$ satisfies some fixed property. To give an example: which modal formulas are first-order definable, given that $R$ is transitive? One of the results is that all modal reduction principles are first-order definable in this case.

I.6 is concerned with the dual question about modally definable $L_0$-formulas. In Kaplan [11] the more general question was asked which classes of models are defined by (sets of) modal formulas. This question was answered in Goldblatt & Thomason [8], using algebraic techniques. For classes of models definable by an $L_0$-sentence their result assumes a very elegant form. This is all we need here, and we give a new proof of the relevant result, which avoids their use of so-called "modal algebras".

In addition a number of preservation results are proved for various model-theoretic notions occurring in Goldblatt & Thomason's theorem. This has the following consequence for modally definable $L_0$-sentences. These are all equivalent to $L_0$-sentences of the form $(\forall x)\phi$, 
where \( \phi \) is an \( L_0 \)-formula with the one free variable \( x \) constructed using atomic formulas, \( \bot \) (a sign standing for a contradiction, the so-called falsum), conjunction, disjunction and restricted quantifiers.

Let us now mention some of the main questions we left open. To begin with, is \( M_1 \) recursively enumerable, and what about \( \bar{M}_1 \)? We doubt if \( M_1 \) and \( \bar{M}_1 \) are even arithmetical, in view of our result (cf. p. 30) that these classes are not provably arithmetical in ZF. Take \( \bar{P}_1 \) to be the class of \( L_0 \)-sentences defined by a modal formula in the global sense. Is \( \bar{P}_1 \) recursively enumerable, and is \( P_1 \) recursive in \( \bar{M}_1 \)? Finally, consider the \( M_1 \) of chapter I.4. It is recursively enumerable, but is it recursive?

Other important questions arise when we consider the notion of completeness, which is not treated in this dissertation. (Proving completeness theorems has been the main activity in modal logic for quite some time.) Consider the class \( \bar{C}_1 \) of modal formulas which are complete with respect to some first-order property of \( R \) expressed by an \( L_0 \)-sentence. It is easy to see that \( \bar{C}_1 \) is arithmetical. K. Fine proved that \( \bar{C}_1 \) is not contained in \( \bar{M}_1 \) (cf. [5]) and S.K. Thomason proved that \( \bar{M}_1 \) is not contained in \( \bar{C}_1 \) (cf. [22]). On the other hand, the modal formulas described in theorem 4.13 are in \( \bar{M}_1 \cap \bar{C}_1 \) (cf. Sahlqvist [16]) and it is an open question if \( M_1 \subset \bar{M}_1 \cap \bar{C}_1 \). What, then, is the exact relation of \( \bar{M}_1 \) to \( \bar{C}_1 \), and can \( \bar{M}_1 \cap \bar{C}_1 \) be characterized in some model-theoretic fashion?

We conclude this introduction with two results about modal formulas. \( L_1 \) is the first-order language with an infinite set of unary predicate constants and one binary predicate constant \( R \). A modal formula as defined above is a formula of the form \( (\forall p_1)\ldots(\forall p_n)\phi(p_1, \ldots, p_n, R) \), where \( \phi \) is an \( m \)-formula as defined below.
1.1 Definition

An $m$-formula is a member of the smallest class $X$ of $L_1$-formulas satisfying

(i) for each unary predicate constant $P$ and each individual variable $x$, $P x \in X$

(ii) if $\alpha \in X$, then $\neg \alpha \in X$

(iii) if $\alpha \in X$ and $\beta \in X$, then $(\alpha \rightarrow \beta) \in X$

(iv) if $\alpha \in X$, then $(\forall y)(Rxy + \alpha) \in X$, provided that $x$ and $y$ are distinct individual variables

In chapter I.2 the traditional $\Box, \Diamond$ -notation is used for modal formulas. These are then translated into formulas of the form

$(\forall P_1 \ldots (\forall P_n) \phi(P_1, \ldots, P_n, R)$, where $\phi$ is an $L_1$-formula of an even more special kind:

1.2 Definition

An $M$-formula is a member of the smallest class $X$ of $L_1$-formulas satisfying

(i) for each unary predicate constant $P$ and each individual variable $x$, $P x \in X$

(ii) if $\alpha \in X$, then $\neg \alpha \in X$

(iii) if $\alpha$ and $\beta$ have the same free variables and are both in $X$, then $(\alpha \rightarrow \beta) \in X$

(iv) if $\alpha \in X$ and $y$ is the free variable of $\alpha$, then $(\forall y)(Rxy + \alpha) \in X$, provided that $x$ is distinct from $y$.

$m$-formulas have at least one free variable, $M$-formulas have exactly one.
1.3 Lemma

Any m-formula $\alpha$ is equivalent to a Boolean combination of M-formulas, each with their free variable among those of $\alpha$.

Proof: The assertion is proved by induction on the complexity of m-formulas. In order to simplify the proof the clauses (iii) and (iv) of the above definitions are temporarily replaced by analogous clauses for conjunction ($\land$), disjunction ($\lor$) and restricted existential quantification ($\exists y)(Rxy \land \ldots$ ). As we are only trying to prove an equivalence this change is harmless.

The cases $\alpha = Px$, $\alpha = \neg \beta$, $\alpha = \beta \land \gamma$ and $\alpha = \beta \lor \gamma$ are trivial.

It remains to consider $\alpha = (\exists y)(Rxy \land \beta)$. By the induction hypothesis $\beta$ is equivalent to a Boolean combination of M-formulas each with their free variable among those of $\beta$. By the theorem on distributive normal forms $\beta$ is then equivalent to a formula of the form $\Sigma_{i=1}^{n} \beta_{ij}$, where $\beta_{ij}$ is an M-formula.

(As for the notation, we stipulate that $\Sigma_{i=1}^{n} \phi_{i} = \text{def} (\phi_{1} \lor \ldots \lor \phi_{n})$ and $\prod_{i=1}^{n} \phi_{i} = \text{def} (\phi_{1} \land \ldots \land \phi_{n})$.)

By standard logic, $(\exists y)(Rxy \land \Sigma_{i=1}^{n} \beta_{ij})$ is equivalent to $\Sigma_{i=1}^{n} (\exists y)(Rxy \land \beta_{ij})$. So it suffices to consider the members of this disjunction. If none of the $\beta_{ij}$'s have a free variable $y$ then $(\exists y)(Rxy \land \prod_{j=1}^{n} \beta_{ij})$ is equivalent to $(\exists y)(Rxy \land (Py \lor \neg Py)) \land \prod_{j=1}^{n} \beta_{ij}$, for an arbitrary unary predicate constant $P$. This is a Boolean combination of M-formulas of the required kind. Otherwise, let $\beta_{1}^{i}$ be the conjunction of those $\beta_{ij}$'s with $y$ as their free variable and let $\beta_{2}^{i}$ be the conjunction of the remainder. Then $(\exists y)(Rxy \land \prod_{j=1}^{n} \beta_{ij})$ is equivalent to $(\exists y)(Rxy \land \beta_{1}^{i} \land \beta_{2}^{i})$, again a Boolean combination of M-formulas of the required kind.

QED.
1.4 Corollary

Any m-formula with one free variable is equivalent to an M-formula.

Proof: A Boolean combination of M-formulas with the same free variable is itself an M-formula. QED.

Before stating the next result we mention a few notational conventions. L₁-models will be denoted by M or N, possibly with subscripts or superscripts. When we want to be explicit we write $M = \langle W, R, V \rangle$, where W is the domain of M, R is the interpretation of the predicate constant R (a harmless autonomy occurs here) and $V(P)$ is the set of those members of W for which $PM$ holds. The sign $\models$, which was used already to denote logical consequence and universal validity, will denote truth in a model when occurring in a context $M \models \phi$. Other model-theoretic notions will be used as well, following the conventions of Chang & Keisler [2]. Two possibly lesser-known notations are used. $FV(\alpha)$ is the set of individual variables occurring free in $\alpha$, and $[t_1/x_1, \ldots, t_n/x_n] \phi$ is the result of simultaneously substituting $t_1$ for $x_1$, $\ldots$, $t_n$ for $x_n$ in $\phi$. More information about terminology is to be found in chapter I.2.

1.5 Definition

$M_1 = \langle W_1, R_1, V_1 \rangle$ is a generated submodel of $M_2 = \langle W_2, R_2, V_2 \rangle$ ($M_1 \subseteq M_2$) if $M_1$ is a submodel of $M_2$ and, for all $w \in W_1$ and $v \in W_2$ such that $R_2 w v$ holds, $v \in W_1$.

1.6 Definition

$\phi$, with the free variables $x_1, \ldots, x_n$, is invariant for generated submodels if, for all models $M_1$ and $M_2$ such that $M_1 \subseteq M_2$ and all
\[ w_1, \ldots, w_n \in W_1, M_1 \models \phi[w_1, \ldots, w_n] \text{ iff } M_2 \models \phi[w_1, \ldots, w_n]. \]

1.7 Definition

C is a \textit{p}-relation between \( M_1 = \langle W_1, R_1, V_1 \rangle \) and \( M_2 = \langle W_2, R_2, V_2 \rangle \) if the following four conditions are satisfied,

(i) the domain of C is \( W_1 \) and the range of C is \( W_2 \)
(ii) for each \( w \in W_1 \) and \( v \in W_2 \) such that \( Cwv \), and each unary predicate constant \( P \), \( w \in V_1(P) \) iff \( v \in V_2(P) \)
(iii) for each \( w, w' \in W_1 \) and \( v \in W_2 \) such that \( R_1ww' \) and \( Cwv \) there exists a \( v' \in W_2 \) with \( R_2vv' \) and \( Cw'v' \)
(iv) for each \( v, v' \in W_2 \) and \( w \in W_1 \) such that \( R_2vv' \) and \( Cwv \) there exists a \( w' \in W_1 \) with \( R_1ww' \) and \( Cw'v' \).

1.8 Definition

\( \phi \), with the free variables \( x_1, \ldots, x_n \), is \textit{invariant for p-relations} if, for all models \( M_1 \) and \( M_2 \), all p-relations C between \( M_1 \) and \( M_2 \), and all \( w_1, \ldots, w_n \in W_1, w'_1, \ldots, w'_n \in W_2 \) such that \( Cw_1w'_1, \ldots, Cw_nw'_n \),

\( M_1 \models \phi[w_1, \ldots, w_n] \text{ iff } M_2 \models \phi[w_1, \ldots, w_n] \).

These concepts are of interest only for formulas with free variables. An \( L_1 \)-sentence invariant for generated submodels is either universally valid or a contradiction, as is easily seen using the methods of chapter I.2.

1.9 Theorem

An \( L_1 \)-formula \( \phi \) containing at least one free variable is equivalent to an \( m \)-formula iff it is invariant for generated submodels and p-relations.
Proof: One direction is easy. Each m-formula is invariant for generated
submodels and p-relations, as a simple induction shows.

On the other hand, let \( \phi \) have this property and let \( \text{FV}(\phi) = \{x_1, \ldots, x_n\} \). Define \( m(\phi) = \{\psi \mid \psi \text{ is an m-formula, } \phi \not\models \psi, \}
\)
\( \text{FV}(\psi) \subseteq \text{FV}(\phi) \). We will show that \( m(\phi) \models \phi \). By the compactness theorem,
this implies \( \psi \models \phi \), for some \( \psi \in m(\phi) \), whence clearly \( \models \phi \leftrightarrow \psi \). Since
the proof uses a construction which recurs at various places in I.6, it
will be given in quite some detail.

Let \( M_1 \models m(\phi)[w_1, \ldots, w_n] \). Introduce individual constants \( w_1, \ldots, w_n \).
The notation \( w \) is consistently used to introduce a unique individual
constant for an object \( w \). Adding \( w_1, \ldots, w_n \) to \( L_1 \) gives a language \( L_{11} \).
\( M_1 \) is then expanded to an \( L_{11} \)-model \( M_{11} \) by interpreting \( w_1 \) as \( w_1, \ldots, w_n \) as
\( w_n \). Let \( \phi^* = [w_1/x_1, \ldots, w_n/x_n]\phi \).

Define \( m(L_{11}) \) to be the class of those sentences of \( L_{11} \) that are
obtained by starting with atomic formulas of the forms P\( x \) or P\( c \) and
applying \( \neg, \land, (\forall y)(Rxy \rightarrow \theta) \) or \( (\forall y)(Rcy \rightarrow \theta \) , where \( x \) and \( y \) are distinct in-
dividual variables and \( c \) is an arbitrary individual constant of \( L_{11} \).
(m-formulas always had at least one free variable, but this relaxation
of the definition generates sentences as well.)

Each finite subset of \( \{\phi^*\} \cup \{\psi \mid \psi \in m(L_{11}) \} \) and \( M_{11} \models \psi \) has a
model. For suppose otherwise. Then, for some \( \psi_1, \ldots, \psi_k \) as described,
\( \phi^* \models \neg(\psi_1 \land \ldots \land \psi_k) \), but, since \( M_1 \models m(\phi)[w_1, \ldots, w_n] \), it follows
that \( M_{11} \models \neg(\psi_1 \land \ldots \land \psi_k) \), contradicting \( M_{11} \models \psi_1 \land \ldots \land \psi_k \). So
there exists a model \( N_{11} \) for the whole set. \( N_{11} \) is an \( L_{11} \)-model
satisfying the following two conditions,

(i) \( N_{11} \models \phi^* \)
(ii) \( N_{11} - m(L_{11}) - M_{11} \),
where (ii) is short for "for each $\phi \in m(L_{11})$, $N_{11} \not\models \phi$ iff $M_{11} \not\models \phi$.

For each $c$ and $w$ such that $c$ is an individual constant in $L_{11}$, $w$ is an element of the domain of $N_{11}$, and $N_{11} \models Rcxw$, add a new constant $k_{cw}$ to $L_{11}$ to obtain $L_{2}$. Then expand $N_{11}$ to an $L_{2}$-model $N_{2}$ by interpreting each $k_{cw}$ as $w$. $m(L_{2})$ is defined in the obvious way.

Each finite subset of $\{\psi \mid \psi \in m(L_{2})$ and $N_{2} \models \psi\}$ $\cup$

$\{Rck_{cw} \mid N_{2} \models Rck_{cw}\}$ has a model which is an expansion of $M_{11}$. To prove this, consider $\psi_{1}, \ldots, \psi_{k}$ as described, together with $Rc_{1}k_{c_{1}w_{1}}, \ldots, Rc_{1}k_{c_{1}w_{1}}$.

Add $Rck_{cw}$ for each $k_{cw}$ occurring in $\psi_{1} \land \ldots \land \psi_{k}$ which is not among $k_{c_{1}w_{1}}, \ldots, k_{c_{w_{1}}}$, say for $k_{c_{1}w_{1}}, \ldots, k_{c_{w_{1}}'}$. Then take distinct variables $x_{1}, \ldots, x_{1}, y_{1}, \ldots, y_{s}$ not occurring in $\psi_{1} \land \ldots \land \psi_{k}$ and substitute them for $k_{c_{1}w_{1}}, \ldots, k_{c_{1}w_{1}}', k_{c_{1}w_{1}}', \ldots, k_{c_{w_{1}}'}$, respectively to obtain

$(\psi_{1} \land \ldots \land \psi_{k})'$. Then $N_{11} \models (\exists x_{1})(Rc_{1}x_{1} \land \ldots \land (\exists x_{1})(Rc_{1}x_{1} \land

(\exists y_{1})(Rc_{1}x_{1} \land \ldots \land (\exists y_{s})(Rc_{s}x_{s} \land (\psi_{1} \land \ldots \land \psi_{k})'))'). This sentence is in $m(L_{11})$ and therefore it also holds in $M_{11}$, since $N_{11} \models m(L_{11}) \models M_{11}$. It is now clear how $M_{11}$ can be expanded to a model for $\{\psi_{1}, \ldots, \psi_{k},

Rc_{1}k_{c_{1}w_{1}}, \ldots, Rc_{1}k_{c_{1}w_{1}}\}$.

Using a well-known model-theoretic argument it follows that the above set has a model $M_{2}$ satisfying the following conditions,

(i) $M_{11} \not\subseteq L_{2}$ (i.e., $M_{11}$ is an $L_{11}$-elementary submodel of $M_{2}$)

(ii) $N_{2} \models m(L_{2}) \models M_{2}$,

where (ii) has the obvious meaning. This situation may be pictured as:

models: $M_{1}, M_{11} \not\subseteq \begin{array}{c}
\begin{array}{c}
\text{11} \text{11} \\
\text{N}_{11}, N_{2}
\end{array}
\end{array} M_{2}$,

languages: $L_{1}, L_{11}, L_{2}, L_{2}$
This construction is repeated, but now starting from $M_2$. For each $c$ and $w$ such that $c$ is a constant in $L_2$, $w$ is an element in the domain of $M_2$ and $M_2 \not\models \text{Rcx}[w]$, add a new constant $k_{cw}$ to $L_2$ to obtain $L_{21}$. $M_2$ is then expanded to an $L_{21}$-model $M_{21}$ by interpreting $k_{cw}$ as $w$. Using an argument similar to the one given above one sees that each finite subset of $\{ \psi \mid \psi \in m(L_{21}) \text{ and } M_{21} \models \psi \} \cup \{ \text{Rck}_{cw} \mid k_{cw} \in L_{21}-L_2 \text{ and } M_{21} \not\models \text{Rck}_{cw} \}$ has a model which is an expansion of $N_2$. Therefore this set has a model $N_{21}$ satisfying the following two conditions,

(i) $N_2 \prec_{L_2} N_{21}$

(ii) $N_{21} - m(L_{21}) - M_{21}$.

In the picture this leads to:

\begin{center}
\begin{tikzpicture}
\node (M1) at (0,0) {$M_1$};
\node (M11) at (1,0) {$M_{11}$};
\node (M2) at (2,0) {$M_2$};
\node (M21) at (3,0) {$M_{21}$};
\node (N11) at (0,-1) {$N_{11}$};
\node (N2) at (2,-1) {$N_2$};
\node (N21) at (3,-1) {$N_{21}$};
\draw (M1) -- (M11) -- (M2); \draw (M2) -- (M21); \draw (N11) -- (N2); \draw (N2) -- (N21);\end{tikzpicture}
\end{center}

languages: $L_1, L_{11}, L_2, L_2, L_{21}$

Iterating this construction yields two elementary chains $M_1, M_2, \ldots$ and $N_{11}, N_{21}, \ldots$ with limits $M$ and $N$, respectively. The required conclusion follows from the assumption on $\phi$ and the fundamental theorem on elementary chains. Since $N_{11} \models \phi^*, N \models \phi^*$. The submodel $N_C$ of $N$ generated by the constants in $\bigcup_n L_n$ is a generated submodel of $N$ and therefore $N_C \models \phi^*$, by the invariance of $\phi$ for generated submodels. The following defines a $p$-relation $C$ between $N_C$ and the generated submodel $M_C$ of $M$ generated by the constants of $\bigcup_n L_n$. Define $\text{Cwv}$ to hold if, for some constant $c \in \bigcup_n L_n$, $w = c^N$ and $v = c^N$. The construction of the
chains guarantees that C satisfies the four properties required. By the
invariance of φ for p-relations, M_C \models \phi^*, and, using the invariance of
φ for generated submodels once more, M \models \phi^*. This implies that M_{11} \models \phi^*,
so M_1 \models \phi[w_1, \ldots, w_n]. \quad \text{QED.}

The use of constants k_{cw}, rather than w, in this proof serves to
avoid the following complication. Let c_1 and c_2 be constants of L_{11} and
let N_2 \models R_{c_1}x[w] and N_2 \models R_{c_2}x[w]. \{R_{c_1}w, R_{c_2}w\} need not have a
model which is an expansion of M_{11}. The method used only leads to the
L_{11}-sentence (\exists x_1)(R_{c_1}x_1 \land R_{c_2}x_1), but this is not a sentence in m(L_{11})
and therefore need not be true in M_{11}. Using k_{c_1w} and k_{c_2w} leads to the
m(L_{11})-sentence (\exists x_1)(R_{c_1}x_1 \land (\exists x_2)R_{c_2}x_2), in which the information
about c_1 and c_2 having a common R-successor is lost.